#78
linda
7/15/03

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants:	Bjork *et al.*	Examiner:	P. Beskar
Serial No.:	09/847,539	Group Art Unit:	1645
Filed:	May 1, 2001		
For:	*"Streptococcal a_2 M Binding Protein"*		

Customer No.: 23719

Kalow & Springut LLP
488 Madison Avenue, 19th Floor
New York, New York 10022

June 18, 2003

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

TRANSMITTAL OF PRIORITY DOCUMENT

Sir:

Enclosed is a copy of the certified priority document for the above-identified application. If any fee is due please charge Deposit Account No. 11-0171 for such sum accordingly.

If there are any questions regarding this matter that need to be resolved, the Examiner is respectfully invited to contact the Applicants' attorney at the telephone number given below. Thank you for your time and attention to this matter.

Respectfully submitted,

Scott D. Locke, Esq.
Registration No.: 44,877
Attorney for Applicant

(212) 813-1600

Certificate of Mailing Under 37 C.F.R. 1.8

I hereby certify that this correspondence is being deposited on the date shown below with the United States Postal Service as first class mail with sufficient postage in an envelope addressed to: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

K Padilla
(Signature)
K Padilla
(Printed Name of Person Signing Certificate)
6/18/03
(Date)

PATENTS • DESIGNS • COPYRIGHT • TRADE MARKS
The Patent Office

INVESTOR IN PEOPLE

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed

Dated 30 April 2003

Patents Form 1/77

Patents Act 1977
(Rule 16)

THE PATENT OFFICE
D
-2 NOV 1998
LONDON

The Patent Office

9
03NOV98 E401829-1 D00192
P01/7700 0.00 - 9823975.9

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form))

The Patent Office

Cardiff Road
Newport
Gwent NP9 1RH

1. Your reference — P.75312 GCW SER

2. Patent application number *(The Patent Office will fill in this part)* — 02 NOV 1998 **9823975.9**

3. Full name, address and postcode of the or of each applicant *(underline all surnames)* — ACTINOVA LIMITED
5 Signet Court,
Swanns Road,
Cambridge, CB5 8LA,
United Kingdom.

Patents ADP number *(if you know it)* — 7415235001

If the applicant is a corporate body, give the country/state of its incorporation — United Kingdom

4. Title of the invention — PROTEIN

5. Name of your agent *(if you have one)* — J A KEMP & CO

"Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)* — 14 SOUTH SQUARE
GRAY'S INN
LONDON WC1R 5LX

26001

Patents ADP number *(if you know it)*

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and *(if you know it)* the or each application number

Country	Priority application number *(if you know it)*	Date of filing *(day / month / year)*

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

Number of earlier application	Date of filing *(day / month / year)*

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? *(Answer "Yes" if:*
a) any applicant named in part 3 is not an inventor, or
b) there is an inventor who is not named as an applicant, or
c) any named applicant is a corporate body:
See note (d)) — Yes

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form	-
Description	41
Claim(s)	2
Abstract	-
Drawing(s)	9

10. If you are also filing any of the following, state how many against each item.

Priority documents	-
Translations of priority documents	-
Statement of inventorship and right to grant of a patent *(Patents Form 7/77)*	-
Request for preliminary examination and search *(Patents Form 9/77)*	-
Request for substantive examination *(Patents Form 10/77)*	-
Any other documents *(please specify)*	-

11. I/We request the grant of a patent on the basis of this application

Signature J.A. Kemp & Co. Date 2 November 1998

12. Name and daytime telephone number of person to contact in the United Kingdom

S.E. ROQUES
0171 405 3292

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

a) If you need help to fill in this form or you have any questions, please contact the Patent Office on 0645 500505.

b) Write your answers in capital letters using black ink or you may type them.

c) If there is not enough space for all the relevant details on any part of this form, please continue of a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.

d) If you have answered "Yes" Patents Form 7/77 will need to be filed.

e) Once you have filled in the form you must remember to sign and date it.

f) For details of the fee and ways to pay please contact the Patent Office.

PROTEIN

Field of the invention

The invention relates to a new family of proteins which are able to bind to α_2-macroglobulin and peptide fragments of this family of proteins.

Background of the invention

Streptococcus pyogenes (group A Streptococcus) is an important human pathogen which causes a variety of diseases such as pharyngitis, impetigo, scarlatina and erysipelas. More severe infections caused by this organism are necrotizing fasciitis and streptococcal toxic shock like syndrome.

S. pyogenes binds several human plasma proteins via its surface proteins. *S. pyogenes* binds to α_2 macroglobulin (α_2M) which is a proteinase inhibitor. α_2M is a glycoprotein of 718 kD composed of two pairs of identical subunits held together by disulphide bonds.

Previous studies have identified a non-proteolytic cell wall protein of 78 kD of Group A Streptococci which binds to α_2M:Chhatwal *et al* J. Bacteriol. (1987) 169(8) 3691-5.

Summary of the invention

The present inventors have identified a new group of proteins which are expressed on the surface of some strains of Group A streptococcus, *S.pyogenes*. These proteins have the ability to bind to α_2-macroglobulin, and show some homology to protein G of Group G streptococcus. The new protein derived from *S.pyogenes* has been termed protein GRAB by the present inventors. The gene encoding this protein is referred to as *grab*.

The present invention relates in particular to a protein which is capable of binding α_2M and which comprises the amino acid sequence of SEQ ID No. 1 or a functional variant thereof. In preferred embodiments, the protein comprises the amino acid sequence of SEQ ID No. 2 or a functional variant thereof, and/or two or more tandem repeats having the amino acid sequence of SEQ ID No 3 or a variant thereof. The protein of the invention may further comprise a cell membrane anchor

region and a hydrophobic transmembrane region. Preferably, the protein consists of the amino acid sequence of any of SEQ ID Nos. 1 to 11 and variants thereof.

The invention also provides:

- a peptide comprising a fragment of at least 6 amino acids in length of a protein having the amino acid sequence of (a) any of SEQ ID Nos 1 to 11 or (b) a variant of any of SEQ ID Nos 1 to 11;
- a DNA sequence which codes for a protein or peptide according to the invention, said DNA sequence being selected from:
  - (a) the DNA sequence of any of SEQ ID Nos 12 to 16 or the complementary strands thereof;
  - (b) DNA sequences which selectively hybridize the DNA sequences defined in (a) or fragments thereof; and
  - (c) DNA sequences which, but for the degeneracy of the genetic code, would hybridize to the DNA sequences defined in (a) or (b) and which sequences code for a protein or peptide having the same amino acid sequence;
- an expression vector comprising a DNA sequence of the invention operably linked to a regulatory sequence;
- a host cell transformed with a DNA sequence of the invention or an expression vector of the invention; and
- a process for producing a protein or peptide of the invention, comprising culturing a host cell of the invention under conditions to provide for expression of the desired protein or peptide.

Description of the figures

Fig. 1 . The binding of radiolabeled α_2M to 10^9 bacteria of different strains of *S. pyogenes* grown to early stationary phase is presented in A (bars represent +SEM, n=3). In B the binding of radiolabeled α_2M to $2x10^8$ KTL3 bacteria was competed with α_2M and with protein G (+/- SD, n=3). In C the scatchard plot for the reaction between α_2M and 10^9 KTL3 bacteria is shown. The shape of the plot suggests two binding sites with different affinities (K_a=$2.0x10^8M^{-1}$ and $5.3x10^6M^{-1}$ respectively).

Fig. 2. A schematic comparison between protein GRAB and protein G is shown in A. The complete nucleotide and amino acid sequence of *grab*/protein GRAB is shown in B.

Fig. 3. Different strains of S. pyogenes were subjected to PCR and the results are set out in (A). From all strains, except from the AP9 strain, a product of between 500 and 850 bp in size could be amplified (A). Schematic comparison of the mature protein GRAB (amino acids 34-188 in Fig 2B) encoded by these strains is shown in B.

Fig. 4. Total RNA from the KTL3 and AP1 strain was isolated from bacteria in early logarithmic phase-EL, late logarithmic phase-LL, early stationary phase-ES, or late stationary phase-LS and was subjected to Northern blotting. A transcript of approximately 600 bp was seen in the KTL3 strain but not in AP1 (left hand panel). The amount of *grab* mRNA was highest in the logarithmic growth phases. In the LS no transcript could be seen. In the right panel the same filter was probed with a probe hybridising to 16S verifying that the same amount of RNA was applied to each well.

Fig. 5. In the left panel of A a commassie stain of an SDS-PAGE is shown where MBP-GRAB, Protein G, and MBP-α chain of β-galactosidase have been separated. The predicted size of MBP-GRAB is 60 kD but it migrates with an apparent size of 80 kDa. The right panel shows an identical SDS-PAGE, blotted to nitrocellucloe and probed with radiolabeled α_2M. In B different amounts of MBP-GRAB was applied to nitrocellulose and probed with radiolabeled α_2M.

Fig. 6. MBP-GRAB was used to inhibit the binding of radiolabeled α_2M to $2x10^8$ KTL3 bacteria. Similarly one synthetic peptide (aa 34-56 in Fig 2B) was able to compete for the binding of α_2M although less efficiency that MBP-GRAB, while an overlapping peptide (aa 51-68 in Fig 2B) did not compete for the binding. Bars represent +/- SD, n=3.

Fig. 7. An internal fragment of *grab*, lacking the part of the gene coding for the cell wall attachment, was cloned into the streptococcal suicide plasmid pFW13 to generate FW-*grab* (see A). pFW-grab was transformed into KTL3 bacteria, to generate MR4. MR4 was completely devoid of α_2M binding as shown in A, (+SD,

n=3). In B the media from overnight cultures of KTL3 and MR4 were precipitated and subjected to SDS-PAGE (left panel) blotted and probed with radiolabeled α_2M (right panel).

Fig. 8. KTL and MR4 bacteria were incubated with α_2M, washed, and bound proteins were eluted. A shows an SDS-PAGE where eluted material from the KTL3 strain, the MR4 strain or trypsin treated KTL3 bacteria (KTL3+T) was separated. As a reference 0.5 μg of α_2M was run on the gel. In B the binding of the radiolabeled fibrinogen was measured after trypsin treatment pf KTL3 or MR4 bacteria. Some bacteria were preincubated with α_2M (+α_2M) and some were not. As can be seen from B, preincubation of KTL3 with α_2M protected the M protein, and thus fibrinogen binding, from trypsin degradation, α_2M pretreatment of MR4 did not affect the fibrinogen binding (+SD n=3).

Fig. 9. In A radiolabeled and activated SCP was mixed with α_2M or plasma and subjected to non reducing SDS PAGE. As references radiolabeled SCP and α_2M were separated on the same gel. Parts of the SCP is seen in a high molecular weight complex with the apparent size of α_2M. In B radiolabeled and activated SCP was added to KTL3 (1), MR4 (3), or the same bacteria preincubated with α_2M (2 and 4 respectively). The binding of SCP was significantly higher to KTL3 bacteria that had been preincubated with α_2M (+SD, n=3). In C the same bacteria that were used in B were resuspended in non-reducing SDS PAGE sample buffer and eluted material was separated by SDS-PAGE. Again radiolabeled SCP and α_2M were separated on the same gel as a reference. From the α_2M pretreated KTL3 bacteria a complex of the size of α_2M could be seen, while in the others only small amounts of SCP was seen.

Detailed description of the invention.

The invention relates generally to proteins which bind α_2M. Binding of α_2M to bacteria or proteins can be determined using radiolabeled α_2M. For example, bacteria can be incubated with radiolabeled α_2M. After centrifugation, radioactivity of the pellets can be determined and expressed as a percentage of added activity over control samples containing no bacteria. The binding of radiolabeled α_2M could also

be competed with non-labeled α_2M or other protein. The Examples below also describe the generation of a mutant strain of *S. pyogenes* which no longer expresses protein GRAB on its surface. This could also be used as a control. Binding of α_2M to proteins can be assessed by immobilizing the proteins on a support such as nitrocellulose and probing with radiolabeled α_2M. After washing, the radioactivity of the bound protein can be determined to give an indication of specific binding of α_2M to bound protein. The Examples below describe one method for evaluation of the binding of α_2M to both bacteria or proteins.

The inventors have identified a region of protein GRAB which can inhibit α_2M binding to *S.pyogenes* which express protein GRAB. The sequence for this region is set out in SEQ ID No.1. The invention relates to proteins comprising the amino acid sequence of SEQ ID No.1 and variants of this sequence. The term variants is used to cover related amino acid sequences which may differ from the exact sequence of SEQ ID No. 1. Variants according to the invention can be identified in a number of different ways as explained in more detail below.

Variant sequences may be identified in protein GRAB produced from other strains of *S. pyogenes.* Partial sequence data for protein GRAB isolated from a number of different strains of *S.pyogenes* is set out in SEQ ID Nos. 7-11. Each of these sequences includes the sequence of SEQ ID No.1 except for a single residue difference in protein GRAB derived from AP1 (SEQ ID No 9). The variation from SEQ ID No.1 is the replacement of isoleucine for threonine at position 18. This sequence is one example of a variant sequence of the invention.

The Examples below show expression of protein GRAB from a number of other strains of *S.pyogenes.* Protein GRAB from these strains may also be used to identify an α_2M binding region or a region which inhibits α_2M binding to protein GRAB expressing *S.pyogenes,* and also to identify sequences which are variants of SEQ ID No.1. The relevant region from such protein GRABs can be identified by alignment of the amino acid sequence data obtained for protein GRAB from other strains with the sequences set out in SEQ ID Nos 1-11. When the maximum alignment is achieved, the relevant region of the protein comprising a variant on SEQ ID No. 1 can readily be identified.

Protein GRAB from other *S.pyogenes* strains can be identified, firstly by investigating the α_2M binding properties of the strain. Subsequently the desired sequence information can be obtained by cloning the genomic DNA and conducting PCR using primers which hybridize to sections of DNA encoding the peptides set out in SEQ ID Nos 1-11. The Examples below demonstrate identification and partial sequencing of protein GRAB derived from a number of *S.pyogenes* strains. In particular, primers hybridizing to the sequences set out in SEQ ID Nos. 17-21 can be used in the cloning and sequencing of protein GRAB from other *S.pyogenes* strains. The region of protein GRAB identified in SEQ ID No. 1 is highly conserved between the different strains of *S.pyogenes*. In general the variant sequences derived from other *S.pyogenes* would be expected to differ by 1, 2, 3, 4, or up to 5 amino acids from SEQ ID No 1, and more likely by 1 or 2 amino acid residues. Proteins having this variant sequence retain the ability to bind to α_2M.

Variants of SEQ ID No.1 also include sequences which vary from SEQ ID No.1 but which are not necessarily derived from naturally occurring protein GRAB. These variants may be described as having a % homology to SEQ ID No.1 or having a number of substitutions within this sequence. Alternatively a variant may be encoded by a polynucleotides which hybridizes to any one of SEQ ID No 12-16, which is discussed in more detail below.

A variant of SEQ ID No. 1 is one which has at least 78 % homology thereto. Preferably the variant will be at least 83 or 87% and more preferably 91 or 96% homologous thereto. Methods of measuring protein homology are well known in the art and it will be well understood by those of skill in the art that in the present context, homology is calculated on the basis of amino acid identity ("hard homology").

Amino acid substitutions may be made, for example from 1, 2 or 3 up to 4, 5 or 6 substitutions in SEQ ID No.1. The modified sequence generally retains the ability to bind α_2M. Conservative substitutions may be made, for example according to the following Table:

ALIPHATIC	Non-polar	G A P
		I L V
	Polar-uncharged	C S T M
		N Q
	Polar-charged	D E
		K R
AROMATIC		H F W Y

Amino acids in the same block in the second column and preferably in the same line in the third column may be substituted for each other.

Preferably, the proteins of the invention comprise an extension to SEQ ID No.1. Thus the protein preferably comprises SEQ ID No.2. The protein may also comprise sequences which are fragments of SEQ ID No.2 which incorporate at least all of SEQ ID No 1. The protein may therefore comprise a sequence of 25 amino acids commencing at the N-terminal of SEQ ID No.2 and may comprise 30, 35, 40, 45 or 50 amino acids of SEQ ID No. 2 up to the entire sequence of 58 amino acids of SEQ ID No 2. The proteins of the invention may also comprise variants of such sequences.

The variants can be defined in a similar manner to the variants if SEQ ID No. 1. Thus the variants may comprise variant sequences derived from other strains of *S.pyogenes*. For example the Examples describe protein GRAB derived from a number of different strains of *S.pyogenes*. SEQ ID Nos. 7-11 set out sequence data for some of these strains. Alignment with SEQ ID No.2 to give the maximum identity in alignment will allow those of skill in the art to determine variant sequences of SEQ ID No. 2.

Other variants can be identified as outlined above from other *S.pyogenes* strains by looking for α_2M binding and cloning and sequencing as before. α_2M binding of variant proteins can be determined by expression cloning and western blotting of the recombinant protein using radiolabeled α_2M.

Variants can also be identified by % homology or have substitutions as

described above. A greater number of substitutions or lower % homology can be tolerated for longer sequences such as larger fragments of SEQ ID No. 2 or the entire sequence. For example, 1, 2, 3 up to about 10 to 15 substitutions in SEQ ID No.2 may be incorporated. Alternatively a variant may have at least 74%, 78% or 81% homology, and preferably has at least 85% or 90%, 95%, 97% or 98% homology. As before the variants preferably maintain the ability to bind α_2M.

The proteins of the invention may also comprise the sequence of SEQ ID No 3 or a variant sequence thereof, or a fragment of either sequence. Preferably the proteins of the present invention further comprise two or more tandem repeats of the sequence SEQ ID No. 3 and variants thereof. The proteins isolated from *S.pyogenes* and termed protein GRAB have at least two repeated sequences adjacent to the C-terminus of SEQ ID No.2 or variant thereof. These repeat sequences have the sequence set out in SEQ ID No.3 or a variant thereof. As can be seen from SEQ ID Nos 7-11, the sequence can show some variation within each repeat both in a single protein GRAB and also between protein GRAB isolated from different strains of *S.pyogenes*. Thus the term repeat as used herein does not mean that an exact repeat of the same sequence is present but simply that a sequence and one or more variants thereof are present, preferably in tandem.

The protein may comprise 2, 3, 4, 5 or 6 or more repeat sequences. Each repeat sequence is generally 28 amino acids in length but may be from 21 up to 35 amino acids in length. Within each protein the length of the repeat sequence therein may vary. For example a protein may comprise a sequence of 28 amino acids followed by a variant repeat sequence of 35 amino acids. The repeat sequence of the invention may adapt a coiled coil structure. This structure is based on hepadic structure of amino acid units which allow the protein to form a coil.

Variants of the repeat sequence of SEQ ID No 3 derived from other strains of *S.pyogenes* can be readily identified by reference to the sequences set out in SEQ ID Nos. 7-11. Each of these sequences has at least two repeats. Repeat sequences derived from protein GRAB from other *S.pyogenes* strains can be identified as outlined above through cloning and sequencing. Other variants encompassed by the present invention are sequences identified by % homology or substitutions as

outlined above for SEQ ID No.1 or Seq ID No. 2. For example a variant may be a repeat having at least 60% homology, preferably at least 70 or 75% up to 85 or 90% up to at least 96% homology with SEQ ID No 3. A variant may have 1, 2 or 3 up to 6, 7, 8 or 9 substitutions in SEQ ID No 3. Preferably the variant retains the heptad structure allowing the region to form a coiled structure. A sequence encoded by a polynucleotide which hybridizes with a polynucleotide encoding a repeat sequence as described herein is also a variant of the invention.

The proteins of the invention may also comprise additional regions such as a cell membrane anchor region and a transmembrane region. The sequence of SEQ ID No.4 comprises a protein having an α_2M binding region, a repeat sequence region and a cell membrane anchor region and transmembrane region. The proteins of the invention can comprise variants of the cell membrane anchor and transmembrane regions as defined above for the other sequences of the protein. Such variants preferably retain the cell membrane anchor function and/or transmembrane function.

It may be desirable to ensure that the transmembrane regions or anchor regions are not present in the protein. For example, if a protein is desired which has the ability to bind α_2M but which will be excreted from the bacterial cell in which it is expressed, the anchor and transmembrane regions are preferably not expressed as part of the protein.

In one preferred embodiment of the present invention, the protein consists essentially of any one of SEQ ID Nos 1-11 and variants thereof as defined above.

The present invention also relates to peptides comprising a fragment of at least 6 amino acids in length of a protein of the invention. In particular, the invention relates to such a peptide comprising a fragment of the protein having the sequence of any one of SEQ ID Nos. 1-11 and variants thereof. Preferably, the fragment will be at least 10, for example at least 12 or 15, amino acids in length. The fragment may be up to 20, 30, 40, 60 or 150 amino acids in length.

In a preferred embodiment, the peptides of the invention have the ability to bind α_2M. This binding can be determined as outlined above. As will be readily appreciated by one skilled in the art, peptides of shorter length preferably comprise a fragment of protein GRAB derived from *S.pyogenes*. For longer peptides, the

sequences may show greater variation as set out above, such as a smaller % homology or greater number of substitutions.

In another embodiment, the peptide comprises a fragment of the repeat sequence or variant thereof, as described above. In this embodiment the peptide may comprise an entire repeat sequence that is of about 28 amino acids in length as outlined above, or two or more repeat sequences in tandem.

Proteins and polypeptides of the invention may be in substantially isolated form. It will be well understood that the proteins or peptides may be mixed with carriers or diluents which will not interfere with the intended purpose of the protein or peptide and still be regarded as substantially isolated. A protein or peptide of the invention may also be in substantially purified form, in which case it will generally comprise the protein or peptide in a preparation in which more than 90%, for example more than 95%, 98% or 99%, by weight of the protein or peptide in the preparation is a protein or peptide of the invention.

Proteins or peptides of the invention may be modified for example by the addition of one or more histidine residues to assist in their identification or purification or by the addition of a signal sequence to promote their secretion from a cell. Some of the signal sequences derived from protein GRAB from a number of *S.pyogenes* strains are set out in SEQ ID Nos. 7-11, and can be seen located N-terminally from the α_2M binding region or SEQ ID No.1 or variant thereof. It may be desirable to provide the peptides or proteins in a form suitable for attachment to a solid support. The proteins or peptides may thus be modified to enhance their binding to a solid support for example by the addition of a cystine residue.

A protein or peptide of the invention may be labelled with a revealing label. The revealing label may be any suitable label which allows the protein or peptide to be detected. Suitable labels include radioisotopes such as ^{125}I,^{35}S or enzymes, antibodies, polynucleotides and linkers such as biotin. Labelled proteins and peptides of the invention may be used in assays for example to assess levels of α_2M. In such assays it may be preferred to provide the peptides attached to a solid support. The present invention also relates to such labelled and/or immobilized protein and peptides packaged in the form of a kit in a container. The kit may optionally contain

other suitable reagent(s), control(s) or instructions and the like.

The proteins of the present invention may be isolated from *S.pyogenes* expressing the protein. Proteins and peptides of the invention may be prepared as fragments of such isolated proteins. The proteins and peptides of the invention may also be made synthetically or by recombinant means. The amino acid sequence of proteins and polypeptides of the invention may be modified to include non-naturally occurring amino acids or to increase the stability of the compound. When the proteins or peptides are produced by synthetic means, such amino acids may be introduced during production. The proteins or peptides may also be modified following either synthetic or recombinant production.

The proteins or peptides of the invention may also be produced using D-amino acids. In such cases the amino acids will be linked in reverse sequence in the C to N orientation. This is conventional in the art for producing such proteins or peptides.

A number of side chain modifications are known in the art and may be made to the side chains of the proteins or peptides of the present invention. Such modifications include, for example, modifications of amino acids by reductive alkylation by reaction with an aldehyde followed by reduction with $NaBH_4$, amidination with methylacetimidate or acylation with acetic anhydride.

The invention also relates to polynucleotides encoding the proteins and peptides of the invention and their use in producing the proteins and peptides of the invention by recombinant means. In particular the invention relates to (a) the DNA sequence of any one of SEQ ID Nos 12 to 16 or the complementary strands thereof; (b) DNA sequences which hybridize to the DNA sequences defined in (a) or fragments thereof; and (c) DNA sequences which, but for the degeneracy of the genetic code, would hybridize to the DNA sequences defined in (a) or (b) and which sequences code for a polypeptide having the same amino acid sequence. Hybridization is typically carried out under conditions of high stringency, such as hybridization buffer of 6x SSC, 0.5% SDS at 65°C. Hybridization conditions equivalent to the conditions described herein could also be used to identify the polynucleotides of the invention.

Polynucleotides of the invention may also comprise corresponding RNA to these DNA sequences. The polynucleotides may be single or double stranded. They may also be polynucleotides which include within them synthetic or modified nucleotides. A number of different types of modification to oligonucleotides are known in the art. These include methylphosphonate and phosphorothioate backbones, addition of acridine or polylysine at the 3' and/or 5' ends of the molecule. For the purposes of the present invention, it is to be understood that the polynucleotides described herein may be modified by any method available in the art.

Preferred polynucleotides of the invention include polynucleotides encoding any of the proteins and peptides described above. Those skilled in the art will understand that numerous different polynucleotides can encode the same protein or peptide as a result of degeneracy of the genetic code.

A nucleotide sequence capable of selectively hybridizing to the DNA sequence of any one of SEQ ID Nos: 12 to 16 or to a DNA sequence complementary to any one of those sequences will be generally at least 70%, preferably at least 80 or 90% and more preferably at least 95% or 97%, homologous to such a DNA sequence. This homology may typically be over a region of at least 20, preferably at least 30, for instance at least 40, 60 or 100 or more contiguous nucleotides of the said DNA sequence.

Any combination of the above mentioned degrees of homology and minimum sized may be used to define polynucleotides of the invention, with the more stringent combinations (i.e. higher homology over longer lengths) being preferred. Thus for example a polynucleotide which is at least 80% homologous over 25, preferably over 30 nucleotides forms one aspect of the invention, as does a polynucleotide which is at least 90% homologous over 40 nucleotides.

Polynucleotides of the invention may be used to produce a primer, e.g. a PCR primer, a primer for an alternative amplification reaction, a probe e.g. labelled with a revealing label by conventional means using radioactive or non-radioactive labels, or the polynucleotides may be cloned into vectors. Such primers, probes and other fragments will be at least 15, preferably at least 20, for example at least 25, 30 or 40 nucleotides in length, and are also encompassed by the term polynucleotides of the

invention as used herein. Examples of primers of the invention are set out in SEQ ID Nos 17 to 21.

Longer polynucleotides will generally be produced using recombinant means, for example using PCR (polymerase chain reaction) cloning techniques. This will involve making a pair of primers (e.g. of about 15-30 nucleotides) to a region of the *grab* which it is desired to clone, bringing the primers into contact with DNA obtained from a bacterial cell, preferably of an *S.pyogenes* strain, performing a polymerase chain reaction under conditions which bring about amplification of the desired region, isolating the amplified fragment (e.g. by purifying the reaction mixture on an agarose gel) and recovering the amplified DNA. The primers may be designed to contain suitable restriction enzyme recognition sites so that the amplified DNA can be cloned into a suitable cloning vector.

Although in general the techniques mentioned herein are well known in the art, reference may be made in particular to Sambrook *et al*, 1989.

Polynucleotides or primers of the invention may carry a revealing label. Suitable labels include radioisotopes such as ^{32}P or ^{35}S, enzyme labels, or other protein labels such as biotin. Such labels may be added to polynucleotides or primers of the invention and may be detected using techniques known *per se*.

Polynucleotides or primers of the invention or fragments thereof labelled or unlabelled may be used by a person skilled in the art in nucleic acid-based tests for detecting or sequencing *grab* in a bacterial sample.

Such tests for detecting generally comprise bringing a bacterial sample containing DNA into contact with a probe comprising a polynucleotide or primer of the invention under hybridizing conditions an detecting any duplex formed between the probe and nucleic acid in the sample. Such detection may be achieved using techniques such as PCR or by immobilizing the probe on a solid support, removing nucleic acid in the sample which is not hybridized to the probe, and then detecting nucleic acid which was hybridized to the probe. Alternatively, the sample nucleic acid may be immobilized on a solid support, and the amount of probe bound to such a support can be detected.

The probes of the invention may conveniently be packaged in the form of a

test kit in a suitable container. In such kits the probe may be bound to a solid support where the assay format for which the kit is designed requires such binding. The kit may also contain suitable reagents for treating the sample to be probed, hybridizing the probe to nucleic acid in the sample, control reagents, instructions, and the like.

Polynucleotides of the invention can be incorporated into a recombinant replicable vector. The vector may be used to replicate the nucleic acid in a compatible host cell. Thus in a further embodiment, the invention provides a method of making polynucleotides of the invention by introducing a polynucleotide of the invention into a replicable vector, introducing the vector into a compatible host cell, and growing the host cell under conditions which bring about the replication of the vector. The vector may be recovered from the host cell. Suitable host cells include bacteria such as *E. coli*, yeast, mammalian cell lines and other eukaryotic cell lines, for example insect cells such as Sf9 cells.

Preferably, a polynucleotide of the invention in a vector is operably linked to a regulatory sequence that is capable of providing for the expression of the coding sequence by the host cell, i.e. the vector is an expression vector. The term "operably linked" refers to a juxtaposition wherein the components described are in a relationship permitting them to function in their intended manner. A regulatory sequence "operably linked" to a coding sequence is ligated in such a way that expression of the coding sequence is achieved under condition compatible with the control sequences.

Such vectors may be transformed or transfected into a suitable host cell as described above to provide for expression of a polypeptide of the invention. This process may comprise culturing a host cell transformed with an expression vector as described above under conditions to provide for expression by the vector of a coding sequence encoding the polypeptides, and optionally recovering the expressed polypeptides.

The vectors may be for example, plasmid or virus vectors provided with an origin or replication, optionally a promoter for the expression of the said polynucleotide and optionally a regulator of the promoter. The vectors may contain one or more selectable marker genes, for example an ampicillin resistance gene in the

case of a bacterial plasmid or a neomycin resistance gene for a mammalian vector. Vectors may be used in vitro, for example for the production of RNA or used to transfect or transform a host cell.

Promoters/enhancers and other expression regulation signals may be selected to be compatible with the host cell for which the expression vector is designed. For example prokaryotic promoters may be used, in particular those suitable for use in *E.coli* strains. When expression of the polypeptides of the invention is carried out in mammalian cells, mammalian promoters may be used. Tissues-specific promoters, for example hepatocyte cell-specific promoters, may also be used. Viral promoters may also be used, for example the Moloney murine leukaemia virus long terminal repeat (MMLV LTR), the promoter rous sarcoma virus (RSV) LTR promoter, the SV40 promoter, the human cytomegalovirus (CMV) IE promoter, herpes simplex virus promoters or adenovirus promoters. All these promoters are readily available in the art.

Vaccines may be prepared from one or more of the proteins or peptides of the invention and a physiologically acceptable carrier or diluent. Typically, such vaccines are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid prior to injection may also be prepared. The preparation may also be emulsified, or the protein encapsulated in a liposome. The active immunogenic ingredient may be mixed with an excipient which is pharmaceutically acceptable and compatible with the active ingredient. Suitable excipients are, for example, water, saline, dextrose, glycerol, ethanol, of the like and combinations thereof.

In addition, if desired, the vaccine may contain minor amounts of auxiliary substances such as wetting or emulsifying agents, pH buffering agents, and/or adjuvants which enhance the effectiveness of the vaccine. Examples of adjuvants which may be effective include but are not limited to: aluminium hydroxide, N-acetyl-muramyl-L-threonyl-D-isoglutamine (thr-MDP), N-acetyl-nor-muramyl-L-alanyl-D-isoglutamine (CGP 11637, referred to as nor-MDP), N-acetylmuramyl-L-alanyl-D-isoglutaminyl-L-alanine-2-(1'-2'-dipalmitoyl-sn-glycero-3-hydroxyphosphoryloxy)-ethylamine (CGP 19835A, referred to as MTP-PE), and

RIBI, which contains three components extracted from bacteria, monophosphoryl lipid A, trehalose dimycolate and cell wall skeleton (MPL+TDM+CWS) in a 2% squalene/Tween 80 emulsion. The effectiveness of an adjuvant may be determined by measuring the amount of antibodies directed against an immunogenic polypeptide containing a GRAB antigenic sequence resulting from administration of this polypeptide in vaccines which are also comprised of the various adjuvants.

The vaccines are conventionally administered parentally, by injection, for example, either subcutaneously or intramuscularly. Additional formulations which are suitable for other modes of administration include suppositories and, in some cases, oral formulations. For suppositories, traditional binders and carriers may include, for example, polyalkylene glycols or triglycerides; such suppositories may be formed from mixtures containing the active ingredient in the range of 0.5% to 10%, preferably 1% to 2%. Oral formulations include such normally employed excipients as, for example, pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharine, cellulose, magnesium carbonate, and the like. These compositions take the form of solutions, suspensions, tablets, pills, capsules, sustained release formulations or powders and contain 10% to 95% of active ingredient, preferably 25% to 70%. Where the vaccine composition is lyophilised, the lyophilised material may be reconstituted prior to administration, e.g. a a suspension. Reconstitution is preferably effected in buffer.

Capsules, tablets and pills for oral administration to a patient may be provided with an enteric coating comprising, for example, Eudragit "S", Eudragit "L", cellulose acetate, cellulose acetate phthalate or hydroxypropylmethyl cellulose.

The proteins or peptides of the invention may be formulated into the vaccine as neutral or salt forms. Pharmaceutically acceptable salts include the acid addition salt (formed with free amino groups of the peptide) and which are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids such as acetic, oxalic, tartaric and maleic. Salts formed with the free carboxyl groups may also be derived from inorganic bases such as, for example, sodium, potassium, ammonium, calcium, or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, 2-ethylamino ethanol, histidine and

procaine.

The vaccines are administered in a manner compatible with the dosage formulation and in such amount will be prophylactically and/or therapeutically effective. The quantity to be administered, which is generally in the range of 5μg to 250μg of antigen per dose, depends on the subject to be treated, capacity of the subject's immune system to synthesize antibodies, and the degree of protection desired. Precise amounts of active ingredient required to be administered may depend on the judgement of the practitioner and may be peculiar to each subject.

The vaccine may be given in a single dose schedule, or preferably in a multiple dose schedule. A multiple does schedule is one in which a primary course of vaccination may be 1-10 separate doses, followed by other doses given at subsequent time intervals required to maintain and or reinforce the immune response, for example at 1 to 4 months for a second dose, and if needed, a subsequent dose(s) after several months. The dosage regimen will also, at least in part, be determined by the need of the individual and be dependent upon the judgement of the practitioner.

The proteins and peptides of the invention which have the ability to bind α_2M may be used to purify α_2M from a sample. Typically, the proteins or peptides of the invention will be bound to a solid support. A sample potentially containing α_2M can be applied to the support to remove α_2M from the sample. If desired, α_2M can then be released from the support for further use.

The proteins and peptides of the invention which are capable of inhibiting binding of α_2M to the surface of streptococci may be used to inhibit such α_2M binding to the bacterial surface. The proteins and peptides can also be used in competition studies to identify other agents which may effect α_2M binding.

The proteins and peptides of the invention can be used to generate antibodies against strains of *S.pyogenes*. The poylnucleotides of the invention can be used in the production of the proteins and peptides of the invention. As outlined above, they may also be used as primers or probes for identification of related genes to *grab*.

Examples

The following Examples illustrate the invention.

Example 1

S. pyogenes bind native α2M via a protein G like protein - Different strains of *S.pyogenes* were tested for their ability to bind radiolabeled native α2M. *S. pyogenes* strains denoted AP are from the Institute of Hygiene and Epidemiology, Prague, Czech Republic. The KTL strains are from the Finnish Institute for health, and the SF370 strain is the ATCC 700294 strain. Bacteria were harvested in early stationary phase or after overnight culture, washed in PBS with 0.05% Tween-20 and 0.02% azide (PBSAT) and resuspended in the same buffer. Concentration of bacteria was determined by spectrophotometry and $2x10^9$ or $4x10^8$ were incubated with radiolabaled α_2M in 225 μl PBSAT for 50 minutes. For competition different amounts of unlabeled inhibitor was added to the tubes. After centrifugation, radioactivity of the pellets was determined and expressed as percentages of the added activity deducing the non-specific binding to the polypropylene tubes.

The results are shown in Fig 1A. The binding ranged from 0-76 % and differed between strains even within a given serotype. No strain bound a trypsin complexed form of α_2M (data not shown).

The KTL3 strain of the clinically important M1 serotype which bound 53% of added α_2M was chosen for further studies. The binding of radiolabeled α_2M to the KTL3 strain could be competed by both non-radioactive α_2M and by protein G from the strain G148, a group G Streptococcus (Fig. 1B). The scatchard plot for the reaction between α_2M and KTL3 bacteria (Fig. 1C) suggests that two different affinities exist, one high affinity interaction $K_a=2.0x10^8M^{-1}$ and one low affinity interaction $K_a=5.3x10^6M^{-1}$. Since the binding of α_2M to the KTL3 strain could be competed by protein G, we used the protein sequence of protein G from G148 in a tBLASTn search against the Streptococcal Genome Sequencing Project database.

A gene coding for a protein with some homology to the α_2M binding E domain of protein G, as well as to the signal sequence and cell-wall attachment of protein G, was identified. The protein was termed protein GRAB from protein G related α_2M binding protein and consisted of 217 amino acids with a deduced molecular weight of 22.8 kDa. In 2A a schematic representation of the homology between protein GRAB and protein G is shown. In Fig 2B the nucleotide and amino

acid sequences are set out. The A region includes the α_2M binding region. Two repeat regions are identified R1 and R2 and are followed by the wall spanning (W) and membrane spanning (M) regions. Protein GRAB was found to contain the consensus sequence for gram-positive surface cell wall anchored proteins (LPXTGX) followed by a stretch of 19 hydrophobic amino acids and a seven residue long hydrophilic C-terminus (Fig. 2B). The first 34 amino acids of protein GRAB showed some homology to the signal sequence (Ss) of protein G and was followed by 35 amino acids with some homology to the E domain of protein G (Fig. 2B). Spacing the regions with homology to protein G two unique repeated regions of 28 amino acids were identified.

Example 2

Distribution of expression of grab - Genomic DNA was prepared from *S. pyogenes*. PCR was performed using *Taq* polymerase (Gibco-BRL, Gaithersburg, MD) and synthetic oligonucleotides hybridizing to *grab*. Primers hybridized to the following nucleotides in figure 2B primer 1: 101-125, primer 2: 101-128, primer 3: 160–185, primer 4: 594-563 and primer 5: 627-605. Restriction enzymes and ligase were from Gibco-BRL and standard ligation, transformation, and plasmid isolation methods were used. For PCR screening and for cloning in pGEM (Promega, Madison, WI) primers 1 and 5 were used. Sequencing of the pGEM-*grab* plasmids was performed using an ABI-470 prism and Taq dyed dideoxy terminator kit (Perkin and Elmer, Norwalk, CT).

The same strains that were used in the screening for α_2M binding were subjected to PCR using primers hybridizing to *grab*. A PCR product could be generated from all strains except for the AP9 strain, but the size of the product varied between 500 base pairs (bp) and 850 bp (Fig. 3A). Sequencing of the PCR product from four strains revealed that the size polymorphism was due to a variable number of 28 amino acids repeats (Fig. 3B). Comparing the sequence from these four strains and the one presented in the Streptococcal Genome Sequencing Project it was found that protein GRAB is highly conserved. Both the C- and N- terminus was nearly 100% conserved while the repeated region showed 86% identity between strains (Fig.

3B). SEQ ID Nos 7 to 11 show partial sequence data for these strains. SEQ ID Nos 12 to 16 show corresponding nucleotide sequences.

The transcription of *grab* was investigated using Northern blotting where total RNA from the KTL3 strain which bound radiolabeled α_2M and a strain that did not (AP1) was electrophorized, blotted, and probed with a PCR product generated from *grab* using primers 1 and 5. Detectable amounts of *grab* RNA was found in KTL3 bacteria but not in AP1 (Fig. 4). The expression was highest in early logarithmic phase and dropped to undetectable amounts in the late stationary phase. The same filters were probed with a probe hybridizing with 16S which verified that the same amount of RNA had been applied to each well (Fig. 4).

Example 3

Protein GRAB binds α_2M via the extreme N-terminus - The DNA encoding the predicted mature protein GRAB (amino acids 34-189 in Fig. 2B) from the KTL3 strain was PCR cloned into the pMal-p2 vector using the EcoR1 and Pst1 sites present in primers 3 and 5 respectively. The vector was transformed into *E.coli*. For molecular cloning purposes the DH5α strain of *Escherichia coli* was used. *E.coli* were grown in Luria Bertoni broth (10g tryptone (Difco), 10g NaCI, and 5g yeast extract (Difco)/l) supplemented with 2 g glucose/l when using the pMal-p2 vector. For growth in petri dishes 15g/l of bacto agar (Difco) was added. When E.coli contained plasmid 100 μg/ml ampicillin (Sigma, St. Louis, MO) was added to the medium. A fusion protein between a maltose binding protein (MBP) and protein GRAB was produced upon induction with IPTG.

The fusion protein was purified by affinity chromatography on an amylase resin, subjected to SDS-PAGE, blotted to a nitrocellulose filter, and the filter was probed with radiolabeled α_2M. Both Protein G and the MBP-GRAB fusion were found to bind α_2M while MBP was unable to bind α_2M (Fig. 5A). Similarly MBP-GRAB, protein G, and MBP were applied in slots to a nitrocellulose membrane and probed with α_2M and it could be concluded that MBP-GRAB bound α_2M while MBP did not (Fig 5B). MBP-GRAB, but not MBP, was found to compete for the binding of radiolabeled α_2M to KTL3 bacteria (Fig 6). Thus both protein GRAB and protein

G can inhibit the binding of α_2M to KTL3 bacteria indicating that the two proteins interact with the same epitope in α_2M. A peptide covering the extreme N-terminus of the mature protein GRAB (amino acids 34-56 Fig. 2B SEQ ID No 1) was synthesized and was able to compete for the binding of α_2M to KTL3 bacteria while an overlapping peptide (amino acids 49-68 in Fig 2B) did not affect the binding (Fig 6). Thus we conclude that the extreme N-terminus of protein GRAB is responsible for binding of α_2M.

Example 4

Generation of a mutant devoid of protein GRAB on its surface - A fragment of *grab* lacking the part encoding the putative cell wall attachment region was generated by PCR from the KTL3 strain using primers 2 and 4. The fragment was cut with XhoI and HindIII which exclusively cut within primers 3 and 4 respectively and cloned into the corresponding site of streptococcal suicide plasmid pFW13 to generate FW-*grab*. This generated a 468 bp internal fragment (nt 113-580 in Fig 2B) of *grab* lacking the part encoding the cell wall attachment (Fig 7A). The plasmid was electroporated into E.coli, plasmid purified and 2μg of pFW-*grab* was electroporated into KTL3 bacteria for homologous recombination (Fig 7A) and several kanamycin resistant transformants were obtained. Using this cloning strategy the mutant should be devoid of surface bound protein GRAB and instead secrete a truncated form (amino acids 34-174 in Fig 2B). One transformant called MR4 was selected and its ability to bind radiolabeled α_2M was completely abolished (Fig 7A).

When the supernatants from an overnight culture of MR4 and KTL3 were precipitated with TCA, subjected to SDS-PAGE, blotted to nitrocellulose, and probed with radiolabeled α_2M it was found that the MR4 strain secreted an α_2M binding protein of 32 kDa which was not found in the KTL3 medium (Fig 7B). The predicted size of the mature protein GRAB is 14.9 kDa, but apparently it migrates much slower in SDS-PAGE which is in concordance with the observation that the MBP-GRAB fusion also migrates slower than predicted. MR4 and KTL3 bacteria had similar growth characteristics in THY medium and the mutant survived as well as the wild type in fresh human blood (data not shown).

Example 5

Hybridization protocol is carried out as follow. Streptococci were grown in Todd-Hewitt broth with 0.2% yeast extract (THY) in 5% CO_2 at 37°C. Genomic DNA was prepared from *S.pyogenes*. 20 μg of DNA was cleaved by *EcoRI* and subjected to agarose gel electrophoresis and capillary blotting (2) onto Hybond-N filters (Amersham, Amersham, UK). A probe was generated by PCR using *Taq* polymerase and synthetic oligonucleotides with sequences GACTCACCTATCGAACAGCCTCG and AGCTTCTTCTGATTGTAAAGCG, hybridising to *grab*. The PCR product was purified on a MicroSpin™ S-200 HR column and was radiolabeled with [α-32P]dATP using bacteriophage T4 polymerase. Membrane was prehybridized in a solution of 6xSSC, 0.5% SDS, 5xDenharts solution, and 100μg/ml salmon sperm DNA at 50°C for two hours. Probe was boiled for five minutes and added to a solution of 6xSSC, 0.5% SDS and 5xDenharts solution and incubated for 14 hours at 65°C. This was followed by washing at room temperature in 2xSSC+0.5% SDS for five minutes and 2xSSC+0.1% SDS for 15 minutes. Further washes were performed in 0.1xSSC+0.5% SDS at 37°C for one hour and in 0.1xSSC+0.1% SDS at 53°C for 30 minutes. Filter was air dried followed by exposure on BAS-III imaging plate and scanning with Bio-Imaging Analyser BAS-2000.

Example 6

α_2M is active and protects the M protein from tryptic digestion when bound to protein GRAB - 10^9 KTL3 or MR4 cells were incubated for 40 minutes with 20 μg α_2M and carefully washed with PBS. Bound α_2M was eluted using 0.1 glycine pH 2 and subjected to SDS-PAGE. In parallel, 0.3μg of trypsin was added to the α_2M treated bacteria and the trypsin was allowed to react with surface bound α_2M for 5 minutes. Free trypsin (not in complex with α_2M) was blocked by adding a fourfold molar excess of SBTI. Cells were pelleted by centrifugation and the resulting pellet was washed once in 1 ml of PBS and resuspended in 150 μl PBS supplemented with 40μg of chloramphenicol/ml. The remaining activity of trypsin in the supernatant and the resuspended pellet was determined using the chromogenic substrate Nα-

bensoyl-L-arginine p-nitroanilide (L-BAPNA) at a concentration of 0.25 mg/ml by measuring OD_{405} after three hours. The obtained value for MR4 was subtracted from that of KTL3 and compared to a standard, where the same assay was run in parallel using purified α_2M of known concentration. For protection assays bacteria were preincubated with α_2M as above, treated with 0.1μg of trypsin in PBS with chloramphenicol as above for 60 minutes at 37°C. Bacteria were diluted 10 times in PBSAT supplemented with 10 mM benzamidine and chloramphenicol as above and $2x10^6$ bacteria were subjected to a binding assay using radiolabeled fibrinogen.

It was found that roughly 0.5μg of α_2M was bound to 10^9 KTL3 bacteria while no band was seen in the eluate from MR4 (Fig 8A). In parallel, the amount of active α_2M bound was estimated by calculating the amounts of α_2M trapped trypsin. This L-BAPNA assay showed that 10^9 KTL3 bacteria bound 0.27+/-0.03 μg of α_2M, which correlates well with what could be eluted from the bacteria (Fig 8A).

The complex between trypsin and α_2M was released from the KTL3 surface since all trypsin activity was found in the supernatant. To determine if this was due to release of th trypsin-α_2M complex from protein GRAB or tryptic degradation of protein GRAB, KTL3 cells were treated with trypsin and SBTI, washed, incubated with α_2M, and bound α_2M was eluted. No α_2M was bound to the trypsin treated cells indicating that protein GRAB was digested by trypsin (Fig 8A). Thus it was concluded that α_2M bound to the surface of KTL3 is active and that protein GRAB is sensitive to trypsin treatment.

A characteristic of *S. pyogenes* M-proteins are their susceptibility to trypsin degradation. This led us to investigate whether preincubation of KTL3 bacteria with α_2M could protect the M protein, and thus fibrinogen binding, from proteolytic degradation by trypsin. It was found that the fibrinogen binding of KTL3 could be preserved by α_2M pretreatment, while the fibrinogen binding of MR4 was unaffected by α_2M pretreatment (Fig 8B).

Example 7

SCP is trapped by α_2M in solution or α_2M bound to S. pyogenes -

Radiolabeled SCP was activated in activation buffer (1 mM EDTA, and 10 mM DTT

in 0.1 M NaAc-HAc, pH 5.0) for 30 minutes at 40°C. Activated SCP (4μl) was mixed with either 4 μg α_2M or 2 μl of plasma in 20 μl PBS, allowed to react for 15 minutes at 37°C, and subjected to SDS-PAGE using non-reducing conditions followed by autoradiography. Alternatively $2x10^9$ bacteria were pretreated with 40 μg α_2M, washed, and incubated with radiolabeled and activated SCP for 15 minutes. Bacteria were pelleted by centrifugation and pellet was washed with 2 ml of PBSAT and recentrifuged. Radioactivity of the pellet was measured and bound material was released by suspension of pellet in non-reducing SDS-PAGE sample buffer. Eluate was subjected to SDS-PAGE and autoradiography.

As outlined above, radiolabeled and activated SCP was mixed with either purified α_2M or with plasma and subjected to non-reducing SDS-PAGE and autoradiography. Part of the radioactivity could be seen as a band with the apparent size of α_2M indicating that a covalent complex had been formed between SCP and α_2M (Fig 9A). Pretreatment of KTL3 and MR4 with α_2M resulted in an increased binding of SCP to KTL3, but not MR4, bacteria (Fig 9B). When bound material was eluted from these bacteria, subjected to SDS-PAGE and autoradiography, it was found that SCP was in complex with α_2M in the case of KTL3, but not in MR4 (Fig 9C). The supernatants were separated on the same gel, and a small proportion of the radioactivity, from the α_2M pretreated KTL3 bacteria, could be seen as band with the apparent size of α_2M (data not shown). Thus we conclude that α_2M in solution or bound to *S. pyogenes* via protein GRAB can trap, and probably also inhibit SCP.

Example 8

Generation of protein GRAB antiserum. The part of protein GRAB encoding aa 34-188 (Fig 2B) was PCR amplified from the KTL3 strain and cloned into the pET11d vector (Pharmacia Biotech, Uppsala, Sweden). Sequencing of the plasmid insert confirmed that the cloned gene was identical to *grab* from the KTL3 strain. Resulting *Escherichia coli* (BL21, Pharmacia Biotech) transformants were grown in 2xYT to OD_{620} of 0.5 and induced using 0.5 mM IPTG. Bacteria were harvested after 3 hours by centrifugation and resuspended in 20mM Tris-HCI pH 8. Bacteria were sonicated and recentrifuged at 8000xg. The bacterial lysate was subjected to ion-

exchange chromatography using a mono Q column and a FPLC system (Pharmacia Biotech). Protein GRAB could be purified to approximately 90% homogeneity.

100μg of protein GRAB, from the ion exchange chromatography, in 500 μl saline was supplemented with 330 μl complete and 170 μl incomplete Freund's adjuvans and material was used to immunize one rabbit. Rabbit was boostered after 6 weeks with 100 μg of protein GRAB in 500 μl saline supplemented 500 μl incomplete Freund's adjuvans. Blood was drawn 2 weeks after boostering and serum was prepared. Serum was used in ELISA experiments where 1 ng of protein GRAB or malose binding protein (MBP, purified from the same strain of E.coli) in 50mM carbonate buffer, pH 9.6 was absorbed to Maxisorb plates (Nunc) at 4°C overnight. Wells were blocked for 1 hour at room temperature using 200μl of PBS+0.05% Tween 20 (PBST), 1% (w/v) BSA (Sigma) and incubated with varying amounts of protein GRAB antiserum or preimmune serum in the same buffer for 2 hours. This was followed by five rounds of washing with PBST and incubation with a peroxidase labelled goat antirabbit antibody (1:3000 in PBST+1 % BSA) for 1 hour at room temperature. After another round of washing 200μl of developing solution (1mg ABTS and 6 mg hydrogen peroxide/ml of Na-citrate pH 4.5) was added to each well and OD_{405} was determined after 20 minutes of incubation at room temperature. Values over 0.3 were regarded as positives. Titer of the preimmune serum was <1:100 and titer of the immune serum was >1:128 000 for protein GRAB and 1:4000 for MBP.

Similarly KTL3 or MR4 bacteria were heat killed at 65°C and 10^8 bacteria were absorbed (as above) to each well. ELISA was performed as above with the exception that protein A (1:5000) was used instead of the secondary antibody. Titer of the preimmune serum was 1:200 for KTL3 and 1:100 for MR4. Titer of the immune serum was 1:4000 for KTL3 and <1:1000 for MR4,

The antiserum was further used for western blotting of a membrane prepared as in Fig 7B. The filter was blocked for 30 minutes at 37°C using PBST with 5% skimmed milk. Immune or preimmune serum was diluted 1:1000 in the blocking buffer and the filter was incubated for 30 minutes at 37°C. The filter was subsequently washed three times for 10 minutes at 37°C using PBST with 0.5M

NaCI. Incubation with a peroxidise labelled goat anti rabbit antibody (1:3000 in blocking buffer) was performed for 30 minutes at 37°C, followed by washing as above. Membranes were incubated with freshly made substrate consisting of 500 μl of 44.4 mM p-Coumaric acid, 100μl 250 mM Luminol (5-amino-2-3-dihydro-1, 4-phtalazinedione), and 6.1 μl of 30% H_2O_2 dissolved in 20ml Tris-HCI pH 8. Membranes were incubated for one minuted at room temperature, dried and put in a plastic bag for exposure on XAR film (Kodak). The preimmune serum showed no reactivity, whereas the immune serum specifically reacted with a band of the same size as the α_2M- binding protein in Fig 7B.

The sequences mentioned herein are set out in the sequence listing below and can be summarised as follows:

SEQ ID No. 1 is the amino acid sequence of positions 34-56 inclusive of strain SF370 as set out in Figure 2B.

SEQ ID No. 2 is the amino acid sequence of positions 34-91 inclusive of strain SF370 as set out in Figure 2B.

SEQ ID No. 3 is the amino acid sequence of positions 92-119 inclusive of strain SF370 as set out in Figure 2B and is one of the repeat sequences of the protein.

SEQ ID No. 4 is the amino acid sequence of positions 34-217 inclusive of strain SF370 as set out in Figure 2B and is the full length mature protein i.e. without the signal sequence.

SEQ ID No. 5 is the amino acid sequence of positions 34-174 inclusive of strain SF370 as set out in Figure 2B. This truncated form of the protein is missing the trans-membrane and wall anchor regions.

SEQ ID No. 6 is the amino acid sequence of positions 34-193 inclusive of strain SF370 as set out in Figure 2B, and does not include the membrane spanning region of the protein.

SEQ ID No. 7 is the amino acid sequence of the full length protein of strain SF370 as set out in Figure 2B including signal sequence.

SEQ ID Nos. 8-11 are partial amino acid sequences for protein GRAB derived from strains KTL9, AP1, AP49 and KTL3 respectively.

SEQ ID Nos. 12-16 are DNA sequences encoding the amino acid sequences of SEQ ID Nos. 7-11 respectively.

SEQ ID Nos. 17-21 are primers derived from SEQ ID No. 12.

SEQUENCE LISTING

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

Val Asp Ser Pro Ile Glu Gln Pro Arg Ile Ile Pro Asn Gly Gly Thr
1 5 10 15

Leu Thr Asn Leu Leu Gly Asn
20

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 58 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Val Asp Ser Pro Ile Glu Gln Pro Arg Ile Ile Pro Asn Gly Gly Thr
1 5 10 15

Leu Thr Asn Leu Leu Gly Asn Ala Pro Glu Lys Leu Ala Leu Arg Asn
20 25 30

Glu Glu Arg Ala Ile Asp Glu Leu Lys Lys Gln Ala Ile Glu Asp Lys
35 40 45

Glu Ala Thr Thr Ala Ile Glu Ala Ala Ser
50 55

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 28 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Ser Asp Ala Leu Glu Ala Leu Ala Asp Gln Thr Asp Ala Leu Gln Ser
1 5 10 15

Glu Glu Ala Ala Val Val Lys Ala Asp Asn Ala Ala
20 25

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 184 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Val Asp Ser Pro Ile Glu Gln Pro Arg Ile Ile Pro Asn Gly Gly Thr
1 5 10 15

Leu Thr Asn Leu Leu Gly Asn Ala Pro Glu Lys Leu Ala Leu Arg Asn
20 25 30

Glu Glu Arg Ala Ile Asp Glu Leu Lys Lys Gln Ala Ile Glu Asp Lys
35 40 45

Glu Ala Thr Thr Ala Ile Glu Ala Ala Ser Ser Asp Ala Leu Glu Ala
50 55 60

Leu Ala Asp Gln Thr Asp Ala Leu Gln Ser Glu Glu Ala Ala Val Val
65 70 75 80

Lys Ala Asp Asn Ala Ala Ser Asp Ala Leu Glu Ala Leu Ala Asp Gln
85 90 95

Thr Asp Ala Leu Gln Ser Glu Glu Ala Glu Val Val Gln Ser Asp Asn
100 105 110

Ala Ala Ser Asp Ala Trp Glu Lys Ala Ala Thr Pro Ile Ala Leu Asp

```
        115                 120                 125

Val Lys Lys Thr Lys Asp Thr Lys Pro Val Val Lys Lys Glu Glu Arg
    130                 135                 140

Gln Asn Val Asn Thr Leu Pro Thr Thr Gly Glu Glu Ser Asn Pro Phe
145                 150                 155                 160

Phe Thr Ala Ala Ala Leu Ala Ile Met Val Ser Thr Gly Val Leu Val
                165                 170                 175

Val Ser Ser Lys Cys Lys Glu Asn
            180
```

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 141 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

```
Val Asp Ser Pro Ile Glu Gln Pro Arg Ile Ile Pro Asn Gly Gly Thr
1               5                   10                  15

Leu Thr Asn Leu Leu Gly Asn Ala Pro Glu Lys Leu Ala Leu Arg Asn
            20                  25                  30

Glu Glu Arg Ala Ile Asp Glu Leu Lys Lys Gln Ala Ile Glu Asp Lys
        35                  40                  45

Glu Ala Thr Thr Ala Ile Glu Ala Ala Ser Ser Asp Ala Leu Glu Ala
    50                  55                  60

Leu Ala Asp Gln Thr Asp Ala Leu Gln Ser Glu Glu Ala Ala Val Val
65                  70                  75                  80

Lys Ala Asp Asn Ala Ala Ser Asp Ala Leu Glu Ala Leu Ala Asp Gln
                85                  90                  95

Thr Asp Ala Leu Gln Ser Glu Glu Ala Glu Val Val Gln Ser Asp Asn
            100                 105                 110

Ala Ala Ser Asp Ala Trp Glu Lys Ala Ala Thr Pro Ile Ala Leu Asp
        115                 120                 125

Val Lys Lys Thr Lys Asp Thr Lys Pro Val Val Lys Lys
    130                 135                 140
```

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 159 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

```
Val Asp Ser Pro Ile Glu Gln Pro Arg Ile Ile Pro Asn Gly Gly Thr
1               5                  10                  15

Leu Thr Asn Leu Leu Gly Asn Ala Pro Glu Lys Leu Ala Leu Arg Asn
            20                  25                  30

Glu Glu Arg Ala Ile Asp Glu Leu Lys Lys Gln Ala Ile Glu Asp Lys
        35                  40                  45

Glu Ala Thr Thr Ala Ile Glu Ala Ala Ser Ser Asp Ala Leu Glu Ala
    50                  55                  60

Leu Ala Asp Gln Thr Asp Ala Leu Gln Ser Glu Glu Ala Ala Val Val
65                  70                  75                  80

Lys Ala Asp Asn Ala Ala Ser Asp Ala Leu Glu Ala Leu Ala Asp Gln
                85                  90                  95

Thr Asp Ala Leu Gln Ser Glu Glu Ala Glu Val Val Gln Ser Asp Asn
            100                 105                 110

Ala Ala Ser Asp Ala Trp Glu Lys Ala Ala Thr Pro Ile Ala Leu Asp
        115                 120                 125

Val Lys Lys Thr Lys Asp Thr Lys Pro Val Val Lys Lys Glu Glu Arg
    130                 135                 140

Gln Asn Val Asn Thr Leu Pro Thr Thr Gly Glu Glu Ser Asn Pro
145                 150                 155
```

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 217 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

Met Gly Lys Glu Ile Lys Val Lys Cys Phe Leu Arg Arg Ser Ala Phe
1 5 10 15

Gly Leu Val Ala Val Ser Ala Ser Val Leu Val Gly Ser Thr Val Ser
20 25 30

Ala Val Asp Ser Pro Ile Glu Gln Pro Arg Ile Ile Pro Asn Gly Gly
35 40 45

Thr Leu Thr Asn Leu Leu Gly Asn Ala Pro Glu Lys Leu Ala Leu Arg
50 55 60

Asn Glu Glu Arg Ala Ile Asp Glu Leu Lys Lys Gln Ala Ile Glu Asp
65 70 75 80

Lys Glu Ala Thr Thr Ala Ile Glu Ala Ala Ser Ser Asp Ala Leu Glu
85 90 95

Ala Leu Ala Asp Gln Thr Asp Ala Leu Gln Ser Glu Glu Ala Ala Val
100 105 110

Val Lys Ala Asp Asn Ala Ala Ser Asp Ala Leu Glu Ala Leu Ala Asp
115 120 125

Gln Thr Asp Ala Leu Gln Ser Glu Glu Ala Glu Val Val Gln Ser Asp
130 135 140

Asn Ala Ala Ser Asp Ala Trp Glu Lys Ala Ala Thr Pro Ile Ala Leu
145 150 155 160

Asp Val Lys Lys Thr Lys Asp Thr Lys Pro Val Val Lys Lys Glu Glu
165 170 175

Arg Gln Asn Val Asn Thr Leu Pro Thr Thr Gly Glu Glu Ser Asn Pro
180 185 190

Phe Phe Thr Ala Ala Ala Leu Ala Ile Met Val Ser Thr Gly Val Leu
195 200 205

Val Val Ser Ser Lys Cys Lys Glu Asn
210 215

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 259 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

```
Ser Ala Phe Gly Leu Val Ala Val Ser Ala Ser Val Leu Val Gly Ser
1               5                  10                  15

Thr Val Ser Ala Val Asp Ser Pro Ile Glu Gln Pro Arg Ile Ile Pro
            20                  25                  30

Asn Gly Gly Thr Leu Thr Asn Leu Leu Gly Asn Ala Pro Glu Lys Leu
        35                  40                  45

Ala Leu Arg Asn Glu Glu Arg Ala Ile Asp Glu Leu Lys Lys Gln Ala
    50                  55                  60

Ile Glu Asp Lys Glu Ala Thr Thr Ala Ile Glu Ala Ala Ser Ser Asp
65                  70                  75                  80

Ala Leu Glu Ala Leu Ala Asp Gln Ala Asp Ala Leu Gln Ser Glu Glu
                85                  90                  95

Ala Ala Val Val Gln Ser Asp Asn Ala Ala Ser Asp Ala Leu Glu Ala
            100                 105                 110

Leu Ala Asp Gln Thr Asp Ala Leu Gln Ser Glu Glu Ala Ala Val Val
        115                 120                 125

Lys Ala Asp Asn Ala Ala Ser Asp Thr Leu Glu Ala Leu Ala Asp Gln
    130                 135                 140

Thr Asp Ala Leu Gln Ser Glu Glu Ala Ala Val Val Lys Ala Asp Asn
145                 150                 155                 160

Ala Ala Ser Asp Thr Leu Glu Ala Leu Ala Asp Gln Thr Asp Ala Leu
                165                 170                 175

Gln Ser Glu Glu Ala Ala Val Val Lys Ala Asp Asn Ala Ala Ser Asp
            180                 185                 190

Thr Leu Glu Ala Leu Ala Asp Gln Thr Asp Ala Leu Gln Ser Glu Glu
        195                 200                 205

Ala Glu Val Val Gln Ser Asp Asn Ala Ala Ser Asp Ala Trp Gly Lys
    210                 215                 220

Ala Ala Thr Pro Ile Ala Leu Asp Val Lys Lys Thr Lys Asp Thr Lys
225                 230                 235                 240

Pro Val Val Lys Lys Glu Glu Arg Gln Asn Val Asn Thr Leu Pro Thr
                245                 250                 255

Thr Gly Glu
```

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 155 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

```
Asp Ser Pro Ile Glu Gln Pro Arg Ile Ile Pro Asn Gly Gly Thr Leu
1               5                  10                  15

Ile Asn Leu Leu Gly Asn Ala Pro Glu Lys Leu Ala Leu Arg Asn Glu
            20                  25                  30

Glu Arg Ala Ile Asp Glu Leu Lys Lys Gln Ala Ile Glu Asp Lys Glu
        35                  40                  45

Ala Thr Thr Ala Ile Glu Ala Ala Ser Ser Asp Ala Leu Glu Ala Leu
    50                  55                  60

Ala Asp Gln Thr Asp Ala Leu Gln Ser Glu Glu Ala Ala Val Val Lys
65                  70                  75                  80

Ala Asp Asn Ala Ala Ser Asp Ala Leu Glu Ala Leu Ala Asp Gln Thr
                85                  90                  95

Asp Ala Leu Gln Ser Glu Glu Ala Glu Val Val Gln Ser Asp Asn Ala
            100                 105                 110

Ala Ser Asp Ala Trp Glu Lys Ala Ala Thr Pro Ile Ala Leu Asp Val
        115                 120                 125

Lys Lys Thr Lys Asp Thr Lys Pro Val Val Lys Lys Glu Glu Arg Gln
    130                 135                 140

Asn Val Asn Thr Leu Pro Thr Thr Gly Glu Glu
145                 150                 155
```

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 271 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

```
Val Ser Ala Val Asp Ser Pro Ile Glu Gln Pro Arg Ile Ile Pro Asn
1               5                  10                  15

Gly Gly Thr Leu Thr Asn Leu Leu Gly Asn Ala Pro Glu Lys Leu Ala
            20                  25                  30

Leu Arg Asn Glu Glu Arg Ala Ile Asp Glu Leu Lys Lys Gln Ala Ile
        35                  40                  45

Glu Asp Lys Glu Ala Thr Thr Ala Ile Glu Ala Ala Ser Ser Asp Ala
    50                  55                  60

Leu Glu Ala Leu Ala Asp Gln Ala Asp Ala Leu Gln Ser Glu Glu Ala
65                  70                  75                  80

Ala Val Val Gln Ser Asp Asn Ala Ala Ser Asp Ala Leu Glu Ala Leu
                85                  90                  95

Ala Asp Gln Ala Asp Ala Leu Gln Ser Glu Glu Ala Ala Val Val Gln
            100                 105                 110

Ser Asp Asn Ala Ala Gly Asp Ala Leu Glu Ala Leu Ala Asp Gln Thr
        115                 120                 125

Asp Ala Leu Gln Ser Glu Glu Ala Ser Val Val Lys Ala Asp Asn Ala
    130                 135                 140

Ala Ser Asp Ala Leu Glu Ala Leu Ala Asp Gln Thr Asp Ala Leu Gln
145                 150                 155                 160

Ser Glu Glu Ala Ser Val Val Lys Ala Asp Asn Ala Ala Ser Asp Ala
                165                 170                 175

Leu Glu Ala Leu Ala Asp Gln Thr Asp Ala Leu Gln Ser Glu Glu Ala
            180                 185                 190

Ala Val Val Lys Ala Asp Asn Ala Ala Ser Asp Ala Leu Glu Ala Leu
        195                 200                 205

Ala Asp Gln Thr Asp Ala Leu Gln Ser Glu Glu Ala Glu Val Val Gln
    210                 215                 220

Ser Asp Asn Ala Ala Ser Asp Ala Trp Glu Lys Ala Ala Thr Pro Ile
225                 230                 235                 240

Ala Leu Asp Val Lys Lys Thr Lys Asp Thr Lys Pro Val Val Lys Lys
                245                 250                 255

Glu Glu Arg Gln Asn Val Asn Thr Leu Pro Thr Thr Gly Glu Glu
            260                 265                 270
```

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 167 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

```
Ala Ser Val Leu Val Gly Ser Thr Val Ser Ala Val Asp Ser Pro Ile
1               5                  10                  15

Glu Gln Pro Arg Ile Ile Pro Asn Gly Gly Thr Leu Thr Asn Leu Leu
            20                  25                  30

Gly Asn Ala Pro Glu Lys Leu Ala Leu Arg Asn Glu Glu Arg Ala Ile
        35                  40                  45

Asp Glu Leu Lys Lys Gln Ala Ile Glu Asp Lys Glu Ala Thr Thr Ala
    50                  55                  60

Ile Glu Ala Ala Ser Ser Asp Ala Leu Glu Ala Leu Ala Asp Gln Thr
65                  70                  75                  80

Asp Ala Leu Gln Ser Glu Glu Ala Ala Val Val Lys Ala Asp Asn Ala
                85                  90                  95

Ala Ser Asp Ala Leu Glu Ala Leu Ala Asp Gln Thr Asp Ala Leu Gln
            100                 105                 110

Ser Glu Glu Ala Glu Val Val Gln Ser Asp Asn Ala Ala Ser Asp Ala
        115                 120                 125

Trp Glu Lys Ala Ala Thr Pro Ile Ala Leu Asp Val Lys Lys Thr Lys
    130                 135                 140

Asp Thr Lys Pro Val Val Lys Lys Glu Glu Arg Gln Asn Val Asn Thr
145                 150                 155                 160

Leu Pro Thr Thr Gly Glu Glu
                165
```

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 654 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

ATGGGAAAAG AAATAAAAGT GAAATGCTTT TTGCGTAGAT CAGCTTTTGG ATTAGTTGCG 60

GTGTCAGCAT CAGTATTAGT CGGTTCAACA GTATCTGCTG TTGACTCACC TATCGAACAG 120

CCTCGAATTA TTCCAAATGG CGGAACCTTA ACTAATCTTC TTGGCAATGC TCCAGAAAAA 180

CTGGCATTAC GTAATGAAGA AAGAGCCATT GATGAATTAA AAAAACAAGC TATTGAGGAT 240

AAAGAAGCTA CGACAGCTAT AGAAGCAGCA AGTTCAGATG CCTTAGAAGC ATTAGCGGAT 300

CAAACAGACG CTTTACAATC AGAAGAAGCT GCGGTTGTTA AAGCGGATAA CGCTGCTAGT 360

GACGCCTTAG AAGCATTGGC GGATCAAACA GACGCTTTAC AATCAGAAGA AGCTGAAGTA 420

GTTCAATCAG ATAACGCTGC TAGTGACGCC TGGGAAAAAG CAGCAACTCC AATCGCTTTA 480

GATGTTAAGA AAACTAAAGA TACAAAACCT GTAGTTAAAA AAGAAGAAAG ACAAAACGTT 540

AATACCCTTC CTACAACTGG TGAAGAGTCT AACCCATTCT TTACAGCTGC TGCGCTTGCA 600

ATAATGGTAA GTACAGGTGT GTTAGTTGTA AGTTCAAAGT GCAAAGAAAA TTAG 654

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 777 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

TCAGCTTTTG GATTAGTTGC GGTGTCAGCA TCAGTATTAG TCGGTTCAAC AGTATCTGCT 60

GTTGACTCAC CTATCGAACA GCCTCGAATT ATTCCAAATG GCGGAACCTT AACTAATCTT 120

CTTGGCAATG CTCCAGAAAA ACTGGCATTA CGTAATGAAG AAAGGGCCAT TGATGAATTA 180

AAAAAACAAG CTATTGAGGA TAAAGAAGCT ACGACAGCTA TAGAAGCAGC AAGTTCAGAT 240

GCCTTAGAAG CATTAGCGGA TCAAGCAGAC GCTTTACAAT CAGAAGAAGC TGCAGTAGTT 300

CAATCAGATA ACGCTGCTAG TGACGCCTTA GAAGCATTGG CGGATCAAAC AGACGCTTTA 360

CAATCAGAAG AAGCTGCGGT TGTTAAAGCG GATAACGCTG CTAGTGACAC TTTAGAAGCA 420

TTGGCGGATC AAACAGACGC TTTACAATCA GAAGAAGCTG CGGTTGTTAA AGCGGATAAC 480

GCTGCTAGTG ACACTTTAGA AGCATTGGCG GATCAAACAG ACGCTTTACA ATCAGAAGAA 540

GCTGCGGTTG TTAAAGCGGA TAACGCTGCT AGTGACACTT TAGAAGCATT GGCGGATCAA 600

ACAGACGCTT TACAATCAGA AGAAGCTGAA GTAGTTCAAT CAGATAACGC TGCTAGTGAC 660

GCCTGGGGAA AAGCAGCAAC TCCAATCGCT TTAGATGTTA AGAAAACTAA AGATACAAAA 720

CCTGTAGTTA AAAAAGAAGA AAGACAAAAC GTTAATACCC TTCCTACAAC TGGTGAA 777

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 469 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

GACTCACCTA TCGAACAGCC TAGAATTATT CCAAATGGCG GAACCTTAAT TAATCTTCTT 60

GGCAATGCTC CAGAAAAACT GGCATTACGT AATGAAGAAA GAGCCATTGA TGAATTAAAA 120

AAACAAGCTA TTGAGGATAA GGAAGCTACG ACAGCTATAG AAGCAGCAAG TTCAGATGCC 180

TTAGAAGCAT TAGCGGATCA AACAGACGCT TTACAATCAG AAGAAGCTGC GGTTGTTAAA 240

GCGGATAACG CTGCTAGTGA CGCCTTAGAA GCATTGGCGG ATCAAACAGA CGCTTTACAA 300

TCAGAAGAAG CTGAAGTAGT TCAATCAGAT AACGCTGCTA GTGACGCCTG GGAAAAAGCA 360

GCAACTCCAA TCGCTTTAGA TGTTAAGAAA ACTAAAGATA CAAAACCTGT AGTTAAAAAA 420

GAAGAAAGAC AAAACGTTAA TACCCTTCCT ACAACTGGTG AAGAGTAAC 469

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 853 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

GTTGCGGTGT CAGCATCAGT ATTAGTCGGT TCAACAGTAT CTGCTGTTGA CTCACCTATC 60

GAACAGCCTC GAATTATTCC AAATGGCGGA ACCTTAACTA ATCTTCTTGG CAATGCTCCA 120

GAAAAACTGG CATTACGTAA TGAAGAAAGA GCCATTGATG AATTAAAAAA ACAAGCTATT 180

GAGGATAAAG AAGCTACGAC AGCTATAGAA GCAGCAAGTT CAGATGCCTT AGAAGCATTA 240

GCGGATCAAG CAGACGCTTT ACAATCAGAA GAAGCTGCAG TAGTTCAATC AGATAACGCT 300

GCTAGTGACG CCTTAGAAGC ATTAGCGGAT CAAGCAGACG CTTTACAATC AGAAGAAGCT 360

GCAGTAGTTC AATCAGATAA CGCTGCTGGT GACGCCTTAG AAGCATTGGC GGATCAAACA 420

GACGCTTTAC AATCAGAAGA AGCTTCGGTT GTTAAAGCGG ATAACGCTGC TAGTGACGCC 480

TTAGAAGCAT TGGCGGATCA AACAGACGCT TTACAATCAG AAGAAGCTTC GGTTGTTAAA 540

GCGGATAACG CTGCTAGTGA CGCCTTAGAA GCATTGGCGG ATCAAACAGA CGCTTTACAA 600

TCAGAAGAAG CTGCGGTTGT TAAAGCGGAT AACGCTGCTA GTGACGCCTT AGAAGCATTG 660

GCGGATCAAA CAGACGCTTT ACAATCAGAA GAAGCTGAAG TAGTTCAATC AGATAACGCT 720

GCTAGTGACG CCTGGGAAAA AGCAGCAACT CCAATCGCTT TAGATGTTAA GAAAACTAAA 780

GATACAAAAC CTGTAGTTAA AAAAGAAGAA AGACAAAACG TTAATACCCT TCCTACAACT 840

GGTGAAGAGT AAC 853

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 504 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

GCATCAGTAT TAGTGGGTTC AACAGTATCT GCTGTGGACT CACCTATCGA ACAGCCTCGA 60

ATTATTCCAA ATGGCGGAAC CTTAACTAAT CTTCTTGGCA ATGCTCCAGA AAAACTGGCA 120

TTACGTAATG AAGAAAGAGC CATTGATGAA TTAAAAAAAC AAGCTATTGA GGATAAAGAA 180

GCTACGACAG CTATAGAAGC AGCAAGTTCA GATGCCTTAG AAGCATTAGC GGATCAAACA 240

GACGCTTTAC AATCAGAAGA AGCTGCGGTT GTTAAAGCGG ATAACGCTGC TAGTGACGCC 300

TTAGAAGCAT TGGCGGATCA AACAGACGCT TTACAATCAG AAGAAGCTGA AGTAGTTCAA 360

TCAGATAACG CTGCTAGTGA CGCCTGGGAA AAAGCAGCAA CTCCAATCGC TTTAGATGTT 420

AAGAAAACTA AAGATACAAA ACCTGTAGTT AAAAAAGAAG AAAGACAAAA CGTTAATACC 480

CTTCCTACAA CTGGTGAAGA GTAA 504

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

AGCTTTTGGA TTAGTTGCGG TGTC 24

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

AGCTTTTGGA TTAGTTGCGG TGTCAGC 27

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

TTGACTCACC TATCGAACAG CCTCG 25

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

AAAACCTGTA GTTAAAAAAG AAGAAAGACA AA 32

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

CCTTCCTACA ACTGGTGAAG AG 22

CLAIMS

1. A protein which is capable of binding to α_2M and which comprises the amino acid sequence of SEQ ID No 1 or a functional variant thereof.

2. A protein according to claim 1 comprising the amino acid sequence of SEQ ID No 2 or a functional variant thereof.

3. A protein according to claim 1 or claim 2 further comprising two or more tandem repeats having the amino acid sequence of SEQ ID No 3 or a variant thereof.

4. A protein according to any one of claims 1, 2 or 3 further comprising a cell membrane anchor region together with a hydrophobic transmembrane region.

5. A protein according to any preceding claim consisting of the amino acid sequence of any of SEQ ID Nos 1 to 11 or a variant thereof.

6. A peptide comprising a fragment of at least 6 amino acids in length of the protein of claim 5.

7. A peptide according to claim 6 comprising a fragment of at least 20 amino acids of the protein of claim 5.

8. A peptide according to claim 6 or 7 which binds α_2M.

9. A peptide according to claim 6 or 7 comprising the acid sequence of SEQ ID NO: 3 or a variant of the said sequence.

10. A peptide according to claim 9 comprising two or more repeats of the amino acid sequence of SEQ ID NO: 3 or of a variant of the said sequence.

11. A DNA sequence which codes for a protein or peptide according to any preceding claim, said DNA sequence being selected from:

(a) the DNA sequence of any of SEQ ID Nos 12 to 16 or the complementary strands thereof;

(b) DNA sequences which selectively hybridize the DNA sequences defined in (a) or fragments thereof; and

(c) DNA sequences which, but for the degeneracy of the genetic code, would hybridize to the DNA sequences defined in (a) or (b) and which sequences code for a protein or peptide having the same amino acid sequence.

12. An expression vector comprising a DNA sequence according to claim 11 operably linked to a regulatory sequence.

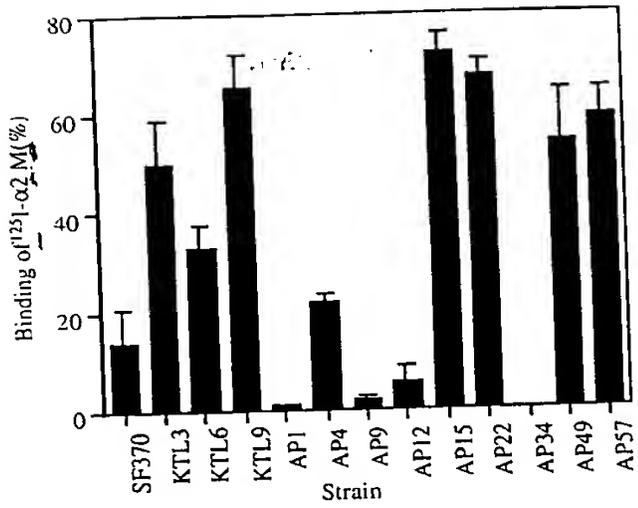
13. A host cell transformed with the DNA sequence of claim 11.

14. A host cell according to claim 13 transformed with the expression vector of claim 12.

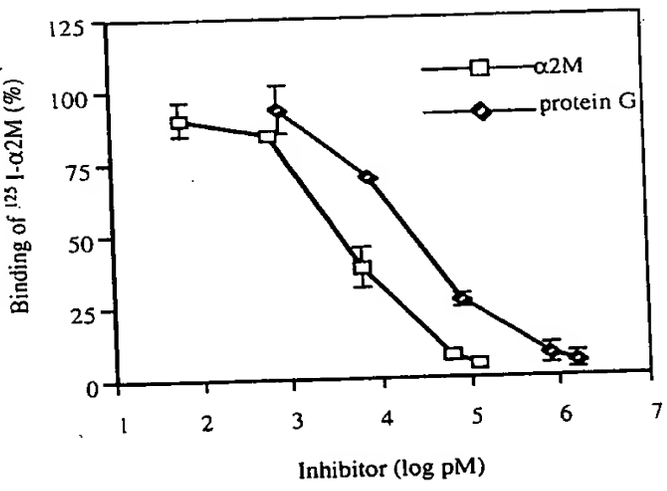
15. A process of producing a protein or peptide according to any of claims 1 to 11, comprising culturing a host cell as defined in claim 13 or 14 under conditions to provide for expression of the desired protein or peptide.

FIG 1

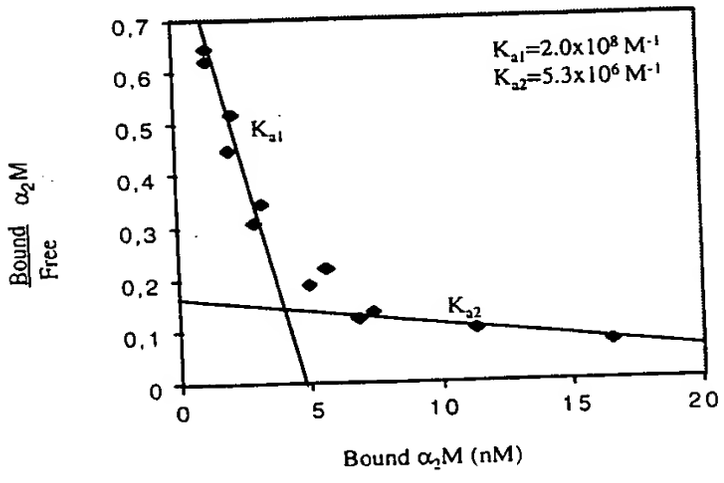
A.


Binding of ^{125}I-α2 M(%)
80
60
40
20
0
SF370
KTL3
KTL6
KTL9
AP1
AP4
AP9
AP12
AP15
AP22
AP34
AP49
AP57
Strain

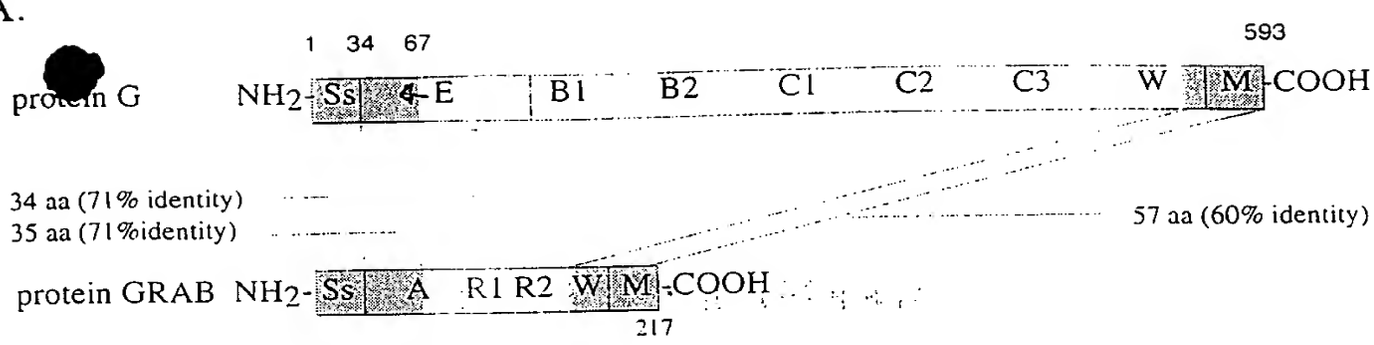
B.


Binding of ^{125}I-α2M (%)
125
100
75
50
25
0
α2M
protein G
1
2
3
4
5
6
7
Inhibitor (log pM)

C.


Bound / Free α2M
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
K_{a1}=2.0x10^8 M^{-1}
K_{a2}=5.3x10^6 M^{-1}
K_{a1}
K_{a2}
0
5
10
15
20
Bound α$_2$M (nM)

A.



B.

```
TTTCCAAATTATCGGTAATTTAATATGCTAATGCATATAAAAATAAAAAAGGAGAAACA      60

ATGGGAAAAGAAATAAAAGTGAAATGCTTTTTGCGTAGATCAGCTTTTGGATTAGTTGCG     120
M  G  K  E  I  K  K  V  K  C  F  L  R  R  S  A  F  G  L  V  A      20
   └──► Ss
GTGTCAGCATCAGTATTAGTCGGTTCAACAGTATCTGCTGTTGACTCACCTATCGAACAG     180
V  S  A  S  V  L  V  G  S  T  V  S  A  V  D  S  P  I  E  Q       40
                                  └──► A
CCTCGAATTATTCCAAATGGCGGAACCTTAACTAATCTTCTTGGCAATGCTCCAGAAAAA     240
P  R  I  I  P  N  G  G  T  L  T  N  L  L  G  N  A  P  E  K       60

CTGGCATTACGTAATGAAGAAAGAGCCATTGATGAATTAAAAAAACAAGCTATTGAGGAT     300
L  A  L  R  N  E  E  R  A  I  D  E  L  K  K  Q  A  I  E  D       80

AAAGAAGCTACGACAGCTATAGAAGCAGCAAGTTCAGATGCCTTAGAAGCATTAGCGGAT     360
K  E  A  T  T  A  I  E  A  A  S  S  D  A  L  E  A  L  A  D      100
                        └──► R1
CAAACAGACGCTTTACAATCAGAAGAAGCTGCGGTTGTTAAAGCGGATAACGCTGCTAGT     420
Q  T  D  A  L  Q  S  E  E  A  A  V  V  K  A  D  N  A  A  S      120
                                                   └──► R2
GACGCCTTAGAAGCATTGGCGGATCAAACAGACGCTTTACAATCAGAAGAAGCTGAAGTA     480
D  A  L  E  A  L  A  D  Q  T  D  A  L  Q  S  E  E  A  E  V      140

GTTCAATCAGATAACGCTGCTAGTGACGCCTGGGAAAAAGCAGCAACTCCAATCGCTTTA     540
V  Q  S  D  N  A  A  S  D  A  W  E  K  A  A  T  P  I  A  L      160
                        └──► W
GATGTTAAGAAAACTAAAGATACAAAACCTGTAGTTAAAAAAGAAGAAAGACAAAACGTT     600
D  V  K  K  T  K  D  T  K  P  V  V  K  K  E  E  R  Q  N  V      180

AATACCCTTCCTACAACTGGTGAAGAGTCTAACCCATTCTTTACAGCTGCTGCGCTTGCA     660
N  T  L  P  T  T  G  E  E  S  N  P  F  F  T  A  A  A  L  A      200
                                  └──► M
ATAATGGTAAGTACAGGTGTGTTAGTTGTAAGTTCAAAGTGCAAAGAAAATTAGTTGCTAT     720
I  M  V  S  T  G  V  L  V  V  S  S  K  C  K  E  N  .            217

TTGTTCTAGCAAATGAAACAAGGGAATCGAAAGATTCTCTTTTT                     764
```

A.

Strain	Serotype	Size of PCR product	Putative number of repeats
AP1	1	500	2[1]
KTL3	1	500	2[1]
KTL6	1	500	2
KTL9	1	650	4[1]
AP4	4	500	2
AP9	9	NA[2]	-
AP12	12	500	2
AP15	15	850	6
AP22	22	650	4
AP34	34	650	4
AP49	49	750	5[1]
AP57	57	500	2

[1]Confirmed by sequencing
[2]Not amplifiable

B.

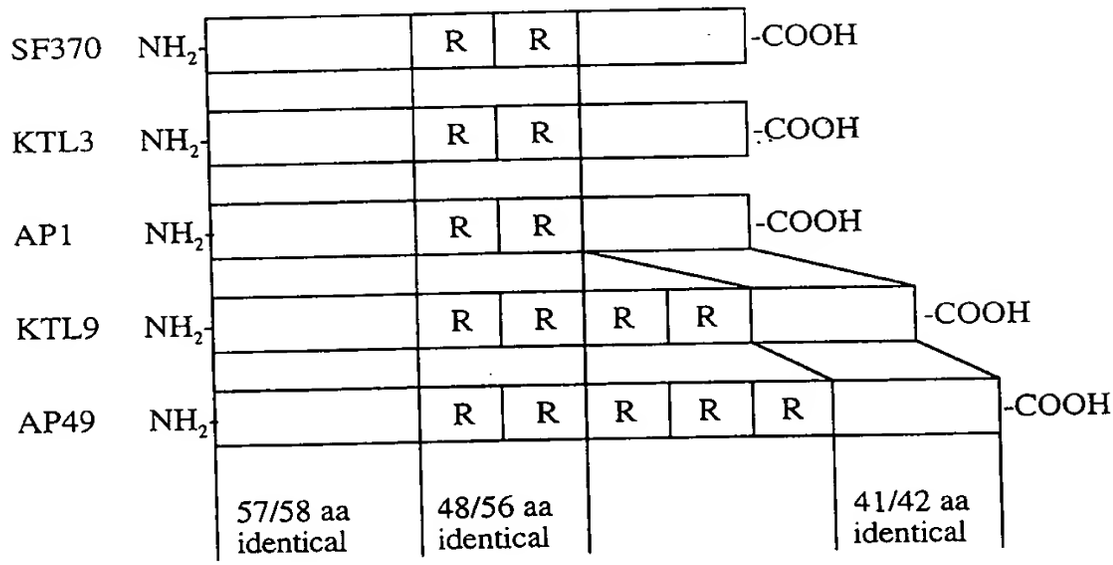
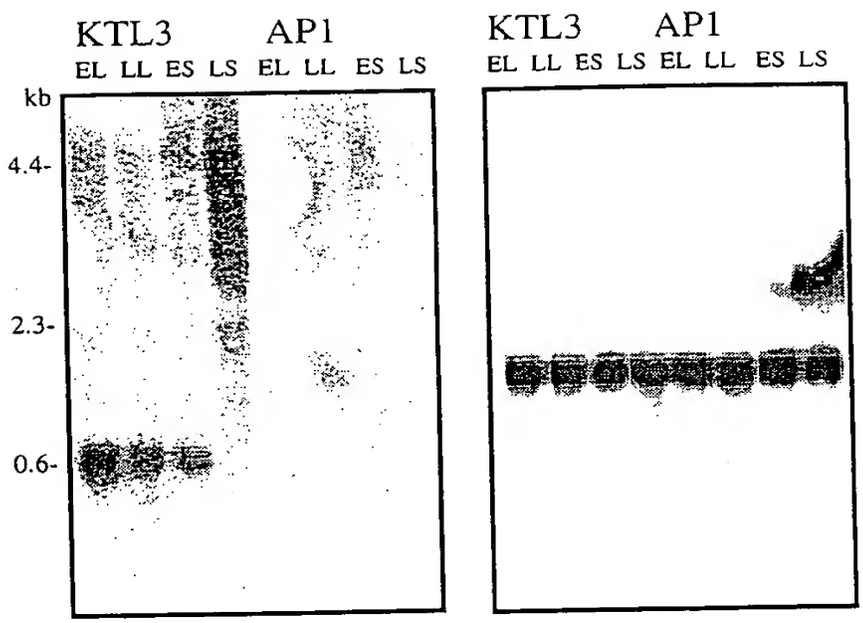


FIG. 4

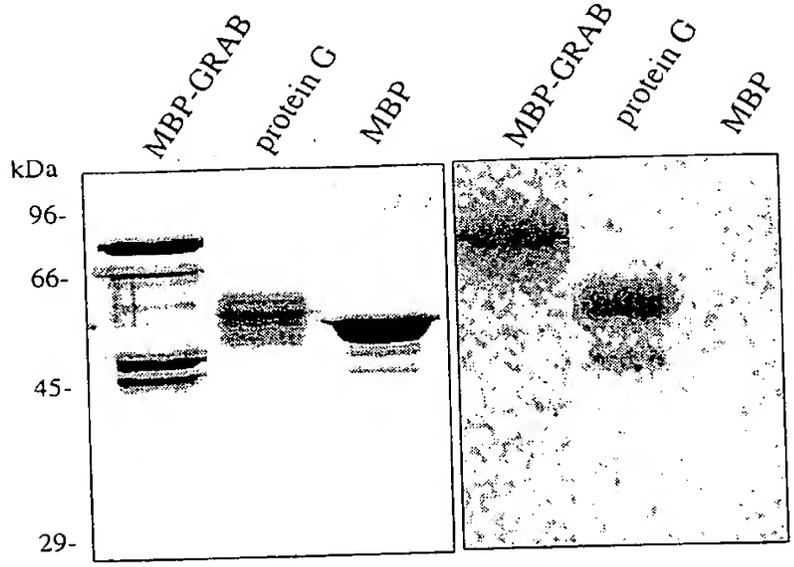


Probe *grab*

Probe 16S

FIG. 5

A.


MBP-GRAB
protein G
MBP
MBP-GRAB
protein G
MBP
kDa
96-
66-
45-
29-

B.

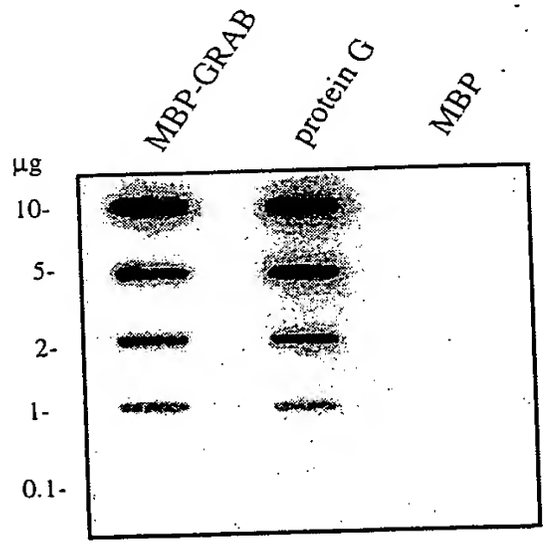
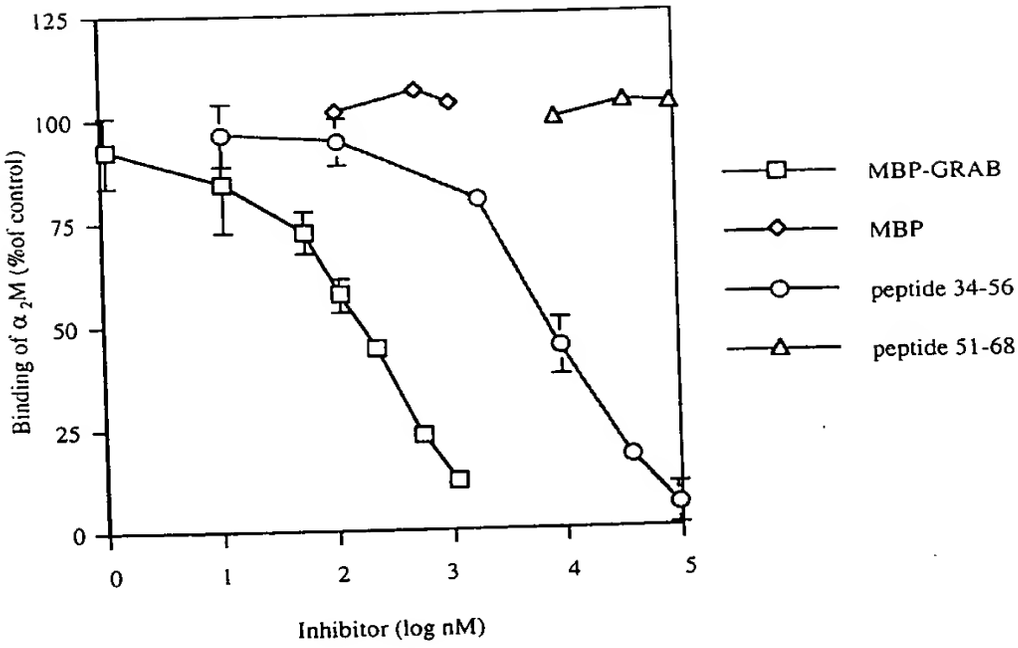
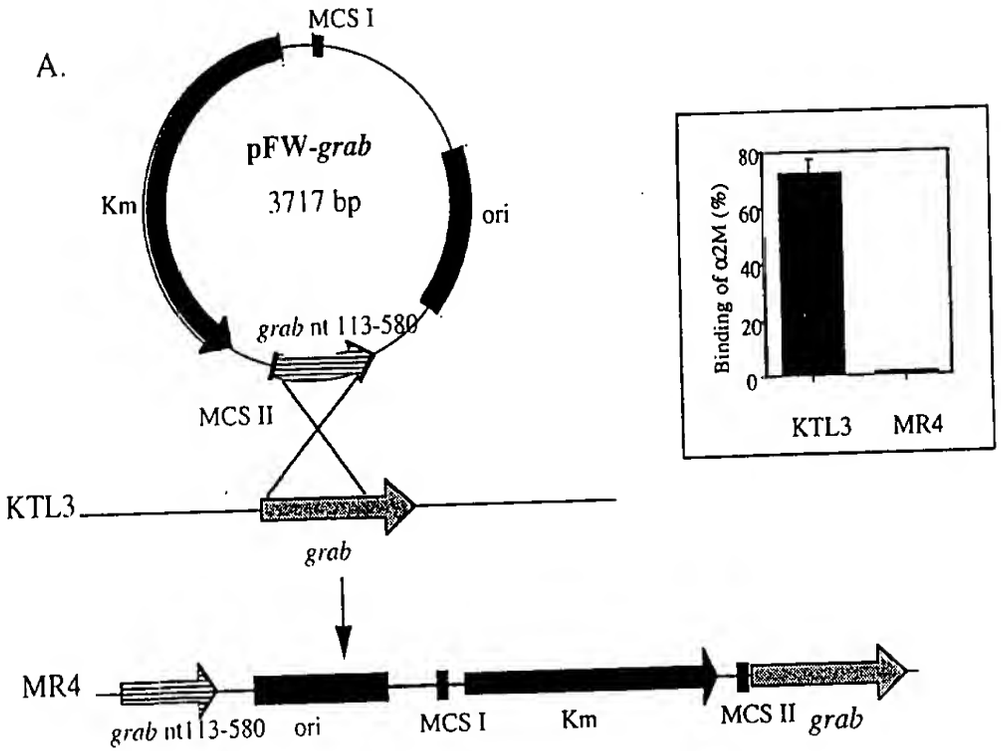
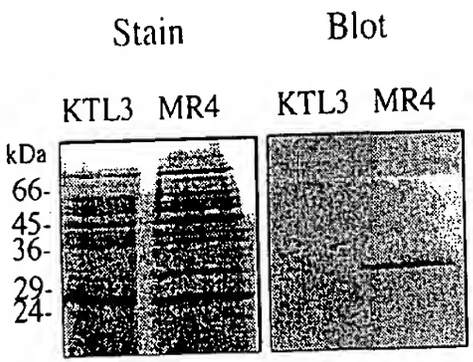

MBP-GRAB
protein G
MBP
µg
10-
5-
2-
1-
0.1-

fig 6 6/9

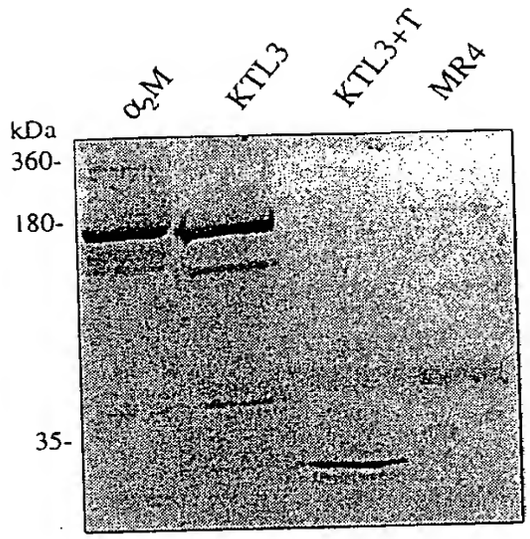

125
100
75
50
25
0
Binding of α_2M (%of control)
0
1
2
3
4
5
Inhibitor (log nM)
MBP-GRAB
MBP
peptide 34-56
peptide 51-68


A.
MCS I
pFW-grab
3717 bp
Km
ori
grab nt 113-580
MCS II
KTL3
grab
MR4
grab nt113-580
ori
MCS I
Km
MCS II
grab
Binding of α2M (%)
80
60
40
20
0
KTL3
MR4

B.


Stain
Blot
KTL3
MR4
KTL3
MR4
kDa
66-
45-
36-
29-
24-

A.


α2M
KTL3
KTL3+T
MR4
kDa
360-
180-
35-

B.

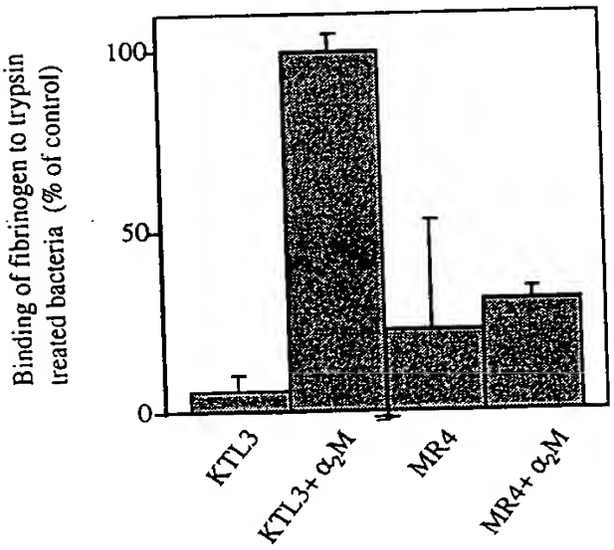
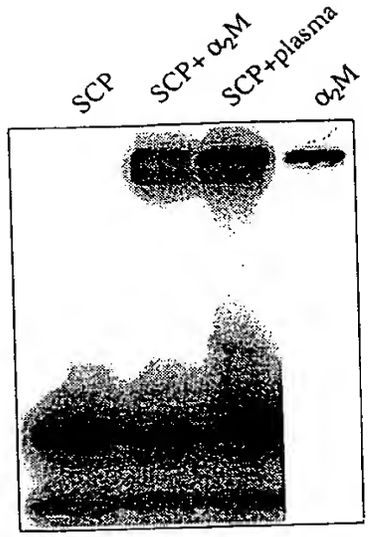
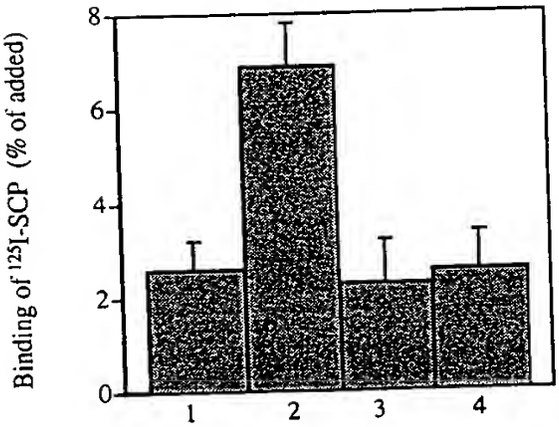

Binding of fibrinogen to trypsin treated bacteria (% of control)
100
50
0
KTL3
KTL3+ α2M
MR4
MR4+ α2M

FIG 9

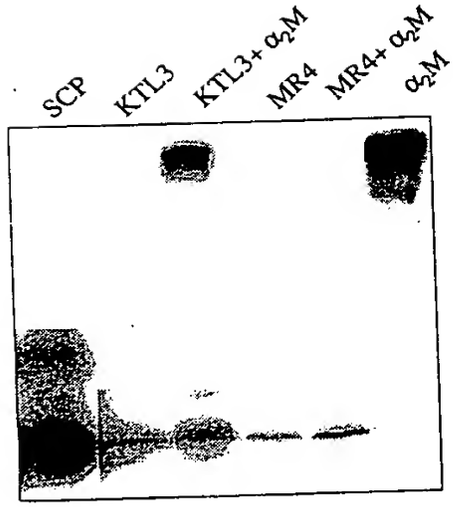
A.


SCP
SCP+ α_2M
SCP+plasma
α_2M

B.


Binding of ^{125}I-SCP (% of added)
8
6
4
2
0
1
2
3
4

C.


SCP
KTL3
KTL3+ α_2M
MR4
MR4+ α_2M
α_2M